184

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Applicants: K. Gordon et al. Examiner: J. Chambers

Serial No.: 07/839,194 Art Unit: 1804

Filed: February 20, 1992

For: TRANSGENIC ANIMALS SECRETING DESIRED PROTEINS INTO MILK

Our Docket No.: IG5-4.4

Corres. and Mail
BOX AF

Hon. Commissioner of Patents
and Trademarks
Washington, DC 20231

95-2206

APPEAL BRIEF

Sir:

This brief is submitted in triplicate in support of Appellant's appeal from the Examiner's final action dated November 7, 1993. A Notice of Appeal was timely mailed to the Patent and Trademark Office in this application on January 27, 1994.

STATUS OF CLAIMS

Claims 1,2,4-9 and 11 are presently pending in this application and are now on appeal.

STATUS OF AMENDMENTS

The claims have not been amended following receipt of the final rejection of February 2, 1994.

SUMMARY OF THE INVENTION

Many proteins which have potential therapeutic applications are relatively scarce in nature, and are therefore extremely expensive to acquire from their native sources. Such sources typically include both plant and animal tissue. With the advent of recombinant technology, such proteins can now be made by genetically modifying a suitable microorganism, such as E. Coli or CHO cells, to express the protein of interest in the culture medium. Recombinant technology has been successful in reducing the relative scarcity and cost of many therapeutic proteins. In addition, recombinant proteins also have the advantage of being relatively free of contaminating biological substances, which makes purification a simpler and less expensive procedure.

While conceptually, the technology for producing recombinant proteins may seem straightforward, there are numerous practical difficulties which must be overcome in order to arrive at a viable production process. Such difficulties include the suitability of the host organism, obtaining significant levels of protein from the culture medium, and ensuring that the expressed protein is biologically active. Even after these difficulties have been overcome, the protein may still be so costly as to make it a marginal product from a purely economic point of view. This has

lead to the active investigation of new approaches to producing proteins of therapeutic interest.

One such approach which has been proposed involves the use of transgenic animals to produce proteins of interest. The present invention resides in DNA constructs which enable the production of proteins in the milk of lactating transgenic animals. The claimed DNA constructs include a gene encoding the protein of interest, DNA of a milk serum protein promoter which does not normally control transcription of the protein, and DNA encoding a peptide enabling secretion of the protein. The milk serum lactalbumin proteins include non-casein milk proteins such as the whey acid protein and α-lactalbumin (page 4, lines 1-15). The secretion signal sequence can be the sequence naturally associated with the desired protein, providing that the protein is one which is normally secreted anyway, or the signal sequence from another secreted protein (page 6, line 13 to page 7, line 8). Other DNA sequences, such as termination sequences, can also be included in the DNA constructs of this invention (page 7, lines 10-15).

The DNA constructs of this invention can be introduced into a mammalian embryo using, for instance, microinjection techniques which are well known in the art (page 8).

ISSUES

I. Whether the invention, as claimed in claims 1,2,5-9 and 11 is adequately enabled by the specification, particularly with respect to milk serum protein promoters.

II. Whether the invention, as claimed in claims 1,2, 4 and 6-9, would have been obvious to one of ordinary skill in the art as of the effective filing date based on Andres et al.

III. Whether the invention, as claimed in claims 5 and 11, would have been obvious to one of ordinary skill in the art as of the effective filing date based on Andres et al., in view of Pennica et al. or Chisari et al.

IV. Whether the invention, as claimed in claims 1,2,4-9 and 11, would have been obvious to one of ordinary skill in the art as of the effective filing date based on Campbell et al., in view of Pennica et al., Chisari et al., Palmiter et al., Ross et al., or Stewart et al.

GROUPING OF CLAIMS

For purposes of this appeal, all of the claims, i.e. claims 1,2,4-9 and 11, stand or fall together.

ARGUMENT

I. The Rejection of Claims 1, 2, 5-9 and 11 Under 35 U.S.C. 112, First Paragraph.

Claims 1, 2, 5-9 and 11 have been finally rejected under 35 U.S.C. 112, first paragraph, as lacking enablement. Claim 4, which is directed to a DNA construct wherein the milk serum promoter is a whey acid promoter, has not been rejected on this basis, and the Examiner has acknowledged that the claim satisfies the enablement requirement.

The Examiner has stated that the specification provides limited guidance on the isolation and identification of regulatory sequences of milk serum proteins other than the whey acid protein promotes. The Examiner asserts that this reasoning is supported by the "unpredictability" of this area of technology, i.e. biotechnology.

In response, appellants assert that there are only a finite number of milk serum protein promoters, and based on the information available as of the effective filing date of this invention are skilled in the art, would have been fully enabled to make and use other such promoters.

A. The Gordon Declaration

The Examiner has criticized the Gordon Declaration as being based on "speculation" and "opinion". Appellants submit that it is improper to dismiss the Gordon Declaration on this basis since the issue to be addressed is enablement, not obviousness. The enablement issue concerns the knowledge of one skilled in the art as of the filing date of the application. This is a hypothetical person in a hypothetical setting, and the issue is whether such a hypothetical person would have possessed sufficient information to

make and use the invention as claimed. This is necessarily a subjective inquiry which is not always amenable to quantification by "concrete supporting evidence". Dr. Gordon is imminently qualified to opine as to the knowledge of one skilled in this art.

More importantly, however, Appellants fail to understand why Dr. Gordon's testimony should not be considered as probative evidence. Her opinion is not based on "thin air", but relies instead on the experiences of a long and extensive career in the young biotechnology industry. In addition, and as will be described in more detail below, Dr. Gordon also bases her opinion on the content of the present specification as well as publications available to the skilled artisan as of the effective filing date of this application. Accordingly, summarily dismissing the Gordon Declaration as based on mere opinions and as lacking supporting evidence is improper.

B. Prior Documentary Evidence

It is Appellants' view that the developed state of the art as of the effective filing date, as shown by published documents, is ample to enable one skilled in the art to practice the invention for milk serum protein promoters. It is important to keep in mind that the Examiner has acknowledged that the specification is fully enabling with respect to the whey acid protein promoter. Thus, the only nonenablement issue to be addressed in this Appeal is the extension of the teachings of the present application to other milk serum protein promoters.

In order to satisfy the enablement requirement, Apellants are entitled to rely on what is disclosed in the specification, as well as disclosures in the prior art, In re Wands, 8 USPQ2d 1400, 1402 (CAFC 1988). As pointed out in the specification, milk proteins are classified as either caseins or milk serum proteins. The milk serum proteins are a defined class of proteins and include, for instance, whey acid protein ("WAP") and alpha-lactalbumin. The promoter is part of the sequence of the protein gene (frequently part of the flanking sequence), and can be obtained using well known techniques once the identity of the gene has been established. In the case of the milk serum proteins, several proteins other than the WAP protein were known prior to the effective filling date of this invention. As an example, the alpha-lactalbumin gene was previously known and characterized by Henninghausen and Sippel in Eur. J. Biochem., 125, 139-140 (1982). A copy of this reference is attached hereto as Attachment 1. This publication was available long before the effective filing date of the present application, and must therefore be considered in any assessment of the enablement issue.

Moreover, the class of molecules encompassed within the definition of milk serum protein promoters is sufficiently circumscribed and delineated so that it would not involve undue experimentation for one skilled in the art to make and use the invention as claimed. See, for instance, In re Wands, 8USPQ2d, supra at 1404, which states that,

> The determination of what constitutes undue experimentation in a given case requires the application of a standard of reasonableness,

> having due regard for the nature of the invention and the state of the art....The test is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed.

Apellants point out that the level of skill in the field of biotechnology is relatively high based on the large proportion of individuals operating within this field who have advanced degrees and specialized training. Such individuals would not require a precise methodology in order to practice a particular invention, but rather would be expected to use a fair amount of experimentation in approaching a task based on this high level of skill. Thus, it would not be unduly burdensome for such individuals to use materials known in the prior art to construct appropriate vectors having a variety of milk serum protein promoters other than the WAP promoter.

C. Burden of Proof

The initial burden of establishing a prima facie case of nonenablement rests on the Examiner. Ex Parte Singh, 17 USPQ2d, 1714, 1715 (CAFC 1990). It is Appellants' position that this burden has not been met in this case. However, assuming, arguendo, that the Examiner has satisfied this burden, Appellants' further submit that they have presented sufficient evidence to

rebut a prima facie case. Such evidence includes both the Declaration of Dr. Gordon and the prior art disclosures of record as more fully discussed above.

Moreover, the holding in Ex parte Singh notwithstanding, the Wands case sets forth eight (8) factors to be considered in determining whether a disclosure is nonenabling. Only one of these factors relates to unpredicatability. Significantly, another factor in the Wands test involves the breadth of the claims, and Appellants maintain that the appealed claims are not overly broad in scope. Other Wands factors include the relative skill of those in the art, which Appellants also maintain is comparatively high. It is not proper to select one of the eight Wands factors in isolation, to the exclusion of other factors which support Appellants' position. The Board's attention is drawn to the fact that In re Wands involved a biotechnology application, and the CAFC concluded that the appealed claims were held to be fully enabling in spite of allegations of unpredictability similar to those presented in the appeal.

D. Equitable Considerations

Finally, affirming this rejection would be tantamount to a holding that each and every transgenic species would have to be actually reduced to practice prior to filing a patent application covering that species. As a practical matter, this would require months, or perhaps years, of additional experimental effort and hundreds of thousands of dollars in added expenses prior to filing a patent application covering such species. Few inventors or small

companies would be able to afford the time or expense to satisfy such a requirement.

Appellants are also concerned that the law in the field of transgenic science is not being uniformly applied in all cases, thus producing anomalous results to the serious disadvantage of those who are diligent in applying for patents, i.e. those who are first to file. This, in effect, rewards those who delay in filing for whatever reason, or who are not particularly diligent in seeking patent protection. Unfortunately, this is precisely the type of abuse the patent system was designed to protect the public against, i.e. to reward those who are not diligent and punish those who are.

As an example, Attachment 2 is a copy of U.S. Patent 4,873,316, which was filed in 1987. A cursory review of this patent reveals that only a single species was actually reduced to practice by the patentee. Yet, the claims of this patent are broadly directed to more than one species. Since the filing date of this patent is less than two years from the effective filing date of the present application, it cannot be seriously contended that transgenic technology advanced so far in that time to justify an entirely different result in the present application. It will be readily apparent that this lack of consistency in the application of settled principles of law works to the serious disadvantage of the most diligent applicants.

The Board is respectfully urged to conclude that there is no proper basis for affirming the rejection of the appealed claims due to a lack of enablement.

II. The Rejection of Claims 1, 2, 4 and 6-9 Under 35 U.S.C. 103.

Claims 1, 2, 4 and 6-9 stand finally rejected as obvious over Andres et al. In rejecting the claims over Andres et al., the Examiner correctly notes that the only difference between the DNA constructs of the present invention and the constructs disclosed in Andres et al. is the presence of a signal peptide secretion sequence in the present claims. However, although signal peptide secretion sequences may have been known in principal prior to the filing date of the present application, Appellants vigorously dispute the proposition that it would have been obvious to include a signal peptide secretion sequence in the DNA constructs of Andres et al., or that the benefit of obtaining secretion of recombinant proteins, such as Ha-ras, would have been expected.

The Andres et al. reference describes the use of the promoter region of the WAP gene in combination with the human Ha-ras gene for expression of the Ha-ras protein in transgenic mice. Andres et al. specifically state that three of the five lines examined, i.e. 60%, did not express the hybrid gene. Moreover, the observed level of expression of the activated Ha-ras gene during lactation was not sufficient to cause transformation of mammary epithelial cells. Finally, Andres et al. state that expression is not limited to the mammary gland, but is also found in brain tissue.

In other words, the attempted expression of the foreign DNA was not successful. Accordingly, one skilled in this art, with the Andres et al. reference before him, would certainly not be motivated to combine the WAP promoter with foreign DNA in order to

obtain satisfactory expression of that DNA in transgenic animals. In this respect, Andres et al. is actually deemed to teach away from the concept of the present invention since it teaches away from the use of the WAP promoter to achieve significant levels of gene expression.

More importantly, however, Andres et al. does not in any way describe or suggest the use of a signal secretion sequence for enabling secretion of the foreign protein into the milk of the host transgenic animal. This is a key aspect of the present invention and, in the absence of such a secretion sequence, the desired protein would remain in the mammary tissue, precluding production of the protein. It is not the objective of the Andres et al. reference to cause secretion of the foreign protein. Rather, Andres et al. merely observe the effect of the expression of the Ha-ras protein in mammary tissue for the purpose of detecting the production of mammary tumors due to the presence of this protein. This is a completely different concept from the concept of the present invention.

The Examiner apparently recognizes the deficiencies of Andres et al. by acknowledging that there is no motivation in Andres et al. to produce a secreted form of the Ha-ras protein. The Examiner then incorrectly concludes that if one skilled in the art desired to secrete recombinant proteins into milk, it would have been obvious to include a signal peptide sequence in the DNA construct. The problem with this reasoning is that Appellants were the first to undertake this approach, i.e. Appellants were the first to appreciate that foreign proteins could be successfully expressed in the milk of a transgenic animal, and

they were the first to propose a tangible approach to achieving this objective. This approach, which ultimately did prove successful, incorporates a signal peptide secretion sequence into a suitable DNA construct, and utilizes this construct to produce a transgenic animal. The transgenic animal was then able to produce a foreign protein (tPA) in its milk.

The Examiner's approach to the question of obviousness is expressly proscribed by the holding in In re O'Farrell, 7 USPQ2d 1673 (CAFC 1988). O'Farrell explicitly holds that "obvious to try" is not the proper standard to apply under 35 USC 103. O'Farrell lists two specific situations where the "obvious to try" approach has been held to be plain error. In one of these situations, "obvious to try" has been held improper when used to explore a new technology where the prior art gave only general guidance as to the particular form of the claimed invention or how to achieve it. In re O'Farrell, supra, at 1681. This is exactly the situation in the present application.

Accordingly, there is no basis for rejecting the appealed claims as obvious on the basis of the Andres et al. reference. Since a prima facie case of obviousness has not been established, for the reasons provided above, a showing of unexpected results is not required.

III. The Rejection of Claims 5 and 11 Under 35 U.S.C. 103.

Claims 5 and 11 have been finally rejected as obvious over Andres et al. in view of Pennica et al. or Chisari et al.

The primary reference, Andres et al., is not applicable to claims 5 and 11 for the reasons discussed above in connection with the previous rejection. Summarizing, Andres et al. is equivocal as to the level of expression achieved, and the significance of this result. In addition, and more significantly, Andres et al. do not in any way teach or suggest the use of a signal peptide sequence in their constructs. Since this is a critical element of the present invention, this omission is fatal.

Pennica et al. describe the use of a plasmid containing the E. coli trp promoter and the cDNA sequence coding for the mature tPA protein for expression of the tPA protein in E. coli. Pennica et al. do not utilize the WAP promoter, and there is no guidance in this reference that would suggest that the disclosed plasmid would be suitable in developing a transgenic animal.

Chisari et al. relate to the development of a transgenic mouse which is capable of expressing the hepatitis B virus surface antigen. As stated in the reference, the expression of the antigen was not tissue specific, and the antigen was detected in the serum of the mouse. Consequently, this reference also fails to describe the use of a DNA sequence for the production of foreign proteins in the milk of a transgenic mammal.

IV. The Rejection of Claims 1,2, 4-9 and 11 Under 35 U.S.C. 103.

Claims 1, 2, 4-9 and 11 have been finally rejected as obvious over Campbell et al. in view of Pennica et al., Chisari et al., Palmiter et al., Ross et al. or Stewart et al. In formulating

this rejection, the Examiner has stated that the combined teachings of the references suggest Appellants' claimed invention, and that the references must be considered together in making this assessment.

Appellants agree with the general proposition that an express suggestion in one prior art reference to be combined with another reference is not necessary for a rejection under 35 USC 103. However, it is also clear that the Patent and Trademark Office has the initial burden under 35 USC 103 of establishing a prima facie case of obviousness, and this burden can only be satisfied by showing some objective teaching in the prior art which would lead to the combination of the relevant teachings of the references. In re Fine, 5 USPQ2d 1596 (Fed. Cir. 1988).

Turning now to Campbell et al., this reference relates to the characterization of the mouse and rat WAP genes. In particular, Campbell et al. describe the 5' and 3' non-coding sequences flanking the coding sequence of the protein. Several potential regulatory sequences of the WAP gene were identified by Campbell et al. which may be related to the regulation of the WAP protein. However, Campbell et al. expressly do not in any way describe or identify the WAP promoter region of the genomic sequence, and there is no appreciation in the reference that the WAP promoter would have any utility for use in a DNA construct to express foreign proteins in the milk of a transgenic animal.

Both the Pennica et al. and Chisari et al. references have been discussed above in connection with the previous ground of rejection. Neither reference is believed to be applicable to the presently claimed invention for reasons discussed previously.

The Palmiter et al. reference discloses tissue-specific expression of recombinant gene products in transgenic animals. This reference is similar in import to Chisari et al., and effectively provides no teaching which would be relevant to the expression of foreign proteins in the milk of transgenic animals.

Similarly, both the Ross et al. and Stewart et al, reference disclose the use of the mouse mammary tumor virus (MTV) promoter, not milk protein regulatory sequences, to drive tissue specific expression of recombinant proteins in transgenic mice. Neither reference describes or suggests the secretion of recombinant proteins into the milk of transgenic animals.

CONCLUSION

The appended claims are believed to be fully enabled and to patentably distinguish over the prior art of record. The rejections are not applicable to the claimed invention for the reasons discussed above.

Accordingly, the Board is respectfully urged to reverse the rejections remaining in this appeal.

Respectfully submitted,

William G. Gosz

William G. Gosz
Attorney for Applicants
Reg. No. 27,787
Phone: (617) 252-7583

APPENDIX

1. A DNA construct containing a gene encoding a protein, said gene being under transcriptional control of a DNA sequence of a mammalian milk serum protein promoter which does not naturally control transcription of said gene, said DNA construct further comprising DNA encoding a peptide enabling secretion of said protein.

2. The DNA construct of claim 1, wherein said secretion-enabling peptide comprises a secretion signal sequence which is cleaved from said secretion protein.

4. The DNA construct of claim 1 wherein said milk serum protein is a whey acid protein.

5. The DNA construct of claim 1 wherein said secretion signal sequence is the secretion signal sequence naturally associated with said protein.

6. The DNA construct of claim 1 wherein said secretion signal sequence is the secretion signal sequence naturally associated with said mammalian milk protein.

7. The DNA construct of claim 1 wherein said DNA construct includes a transcriptional stop sequence.

8. The DNA construct of claim 7 wherein said stop sequence is derived from SV40 virus DNA.

9. The DNA construct of claim 7 wherein said stop sequence is contained in the polyadenylation sequence of SV40.

11. The DNA construct of claim 1 wherein said protein is human tissue plasminogen activator or hepatitis B surface antigen.

ATTACHMENT 1

34. Lüderitz, O., Galanos, C., Lehmann, V., Nurminen, M., Rietschel, E. Th., Rosenfelder, G., Simmon, M. & Westphal, O. (1973) *J. Infect. Dis. 128*, S17–S19.
35. Chiller, J. M., Skidmore, B. J., Morrison, D. C. & Weigle, W. O. (1973) *Proc. Natl Acad. Sci. USA. 70*, 2129–2133.
36. Rosenstreich, D. L., Nowotny, A., Chused, T. & Mergenhagen, S. E. (1973) *Infect. Immun. 8*, 406–411.
37. Passwell, J. H., Dayer, J. M., Gass, K. & Edelson, P. J. (1980) *J. Immunol. 125*, 910–913.
38. Bona, C. A. (1973) *J. Infect. Dis. 128*, S72–S81.
39. Dayer, J. M., Stephenson, M. L., Schmidt, E., Karge, W. & Krane, S. M. (1981) *FEBS Lett. 124*, 253–257.
40. Dayer, J. M., Krane, S. M. & Mergenhagen, S. E. (1981) *Proc. Natl Acad. Sci. USA. 78*, 2474–2477.
41. Dayer, J. M., Russel, R. G. G. & Krane, S. M. (1977) *Science (Wash. DC) 195*, 181–183.
42. Dayer, J. M., Robinson, D. R. & Krane, S. M. (1977) *J. Exp. Med. 145*, 1399–1404.
43. Wahl, L. M., Olsen, C. E., Sandberg, A. L. & Mergenhagen, S. E. (1977) *Proc. Natl Acad. Sci. USA. 74*, 4955–4958.
44. Dahlgren, M. E., Davies, P. & Bonney, R. J. (1980) *Biochim. Biophys. Acta. 630*, 338–351.
45. Samuelsson, B., Goldyne, M., Granström, E., Hamarström, S. & Malmsten, C. (1978) *Annu. Rev. Biochem. 47*, 997–1039.
46. Kuehl, F. A., Cirillo, V. J. & Olen, H. G. (1976) in *Prostaglandins: Chemical and Biochemical Aspect* (Karim, S. M., ed.) pp. 191–225, MTP Press, Lancaster.
47. Brønstad, G. O. & Christoffersen, T. (1981) *Eur. J. Biochem. 117*, 369–374.
48. McCarthy, J. B., Wahl, S. M., Rees, J. C., Olsen, C. E., Sandberg, A. L. & Wahl, L. M. (1980) *J. Immunol. 124*, 2405–2409.
49. Hsueh, W., Crussi, G. & Hanneman, E. (1980) *Nature (Lond.) 283*, 80–82.
50. Sadowski, S. J., Kuehl, F. A., Jr & Davies, P. (1977) *Nature (Lond.) 269*, 149–151.
51. Rouzer, C. A., Scott, W. A., Kempe, J. & Cohen, Z. A. (1980) *Proc. Natl Acad. Sci. USA. 77*, 4279–4282.
52. Rietschel, E. Th., Schade, U., Lüderitz, O., Fischer, H. & Peskar, B. A. (1980) in *Microbiology 1980* (Schlessinger, D., ed.) pp. 66–72, Am. Soc. Microbiol., Washington, DC.
53. Schade, U. & Rietschel, E. Th. (1980) in *Natural Toxins* (Eaker, D. & Wadström, T., eds) pp. 271–277, Pergamon Press, Oxford and New York.

R. Bhatnagar and K. Decker*, Biochemisches Institut der Medizinischen Fakultät der Albert-Ludwigs-Universität Freiburg, Hermann-Herder-Straße 7, D-7800 Freiburg, Federal Republic of Germany

U. Schade and E. Th. Rietschel, Forschungsinstitut Borstel, Parkallee 1, D-2061 Borstel, Federal Republic of Germany

* To whom correspondence should be addressed.

Eur. J. Biochem. 125, 131–141 (1982)
© FEBS 1982

Characterization and Cloning of the mRNAs Specific for the Lactating Mouse Mammary Gland

Lothar G. HENNIGHAUSEN and Albrecht E. SIPPEL

Institut für Genetik, Universität zu Köln, Köln

(Received December 24, 1981/March 4, 1982)

We have characterized and cloned the lactation-specific mRNAs of mouse mammary glands. The group of eight milk-protein-specific mRNAs were identified (a) by size and antigenic properties of their translation products *in vitro* and (b) by characterization of their respective cDNA clones.

Two α-caseins (43 kDa and 39 kDa) are encoded by mRNAs of 1600 nucleotides and two β-caseins (26 kDa) are encoded by mRNAs of 1450 nucleotides in length. Three smaller caseins, γ-casein (23.7 kDa), δ-casein (21 kDa) and κ-casein (14.5 kDa) are synthesized by mRNAs of 880, 1150 and 860 nucleotides. Beside these casein mRNAs a mammary specific 620 nucleotide mRNA codes for a novel acidic whey protein (13.7 kDa).

cDNA clones corresponding to the mRNAs for the lactation-specific proteins have been isolated from a mammary-specific cDNA library. Cloned α-casein cDNA hybridizes to both α-casein specific mRNAs and cloned β-casein cDNAs hybridize with both β-casein specific mRNAs. By RNA blot analysis we show that the cloned cDNAs for mouse α-casein, β-casein, γ-casein and κ-casein and the acidic whey protein cross-hybridize with mRNAs of the rat, demonstrating partial sequence homology between the corresponding mRNAs of those species.

To learn more about how genes are regulated it was fruitful in recent years to study systems, like the chicken oviduct, in which the expression of several genes are under coordinated control of hormones [1]. The mammary gland provides a system where the interactions of several steroid and peptide hormones on the expression of milk protein genes can be studied [2]. After development and differentiation of the mammary gland, regulated by the steroid and peptide hormones [3], the synthesis of milk proteins, in the mature, lactating tissue is stimulated by glucocorticoids and prolactin [4] and suppressed by progesterone [5]. In recent years several strategies have been applied to correlate milk gene expression to the hormonal and differentiated state of the mammary gland. First, studies on mammary gland tissue from animals during different stages of gestation and lactation [6–8] demonstrated the appearance of milk-protein-specific mRNAs by mid pregnancy in rat and in rabbit. In the guinea-pig α-lactalbumin transcripts appeared late in pregnancy, whereas casein-specific mRNA was detectable only after parturition [9]. In a second approach the effect of steroid and peptide hormones on casein mRNA synthesis in the mammary gland of pseudopregnant animals was investigated [10, 11]. Third, murine mammary gland organ cultures have been employed to study the mechanism by which peptide hormones and steroid hormones regulate gene expression [4, 12]. From these studies it was deduced that the hormones regulate transcriptional activity of the milk-protein-specific genes as well as the stability of the respective mRNAs. Specific RNA concentrations hitherto were measured by translation *in vitro* or mRNA titration with cDNA probes derived from purified mRNA. In order to increase specificity of hybridization probes we and others [13–15] started to clone mammary-gland-specific cDNAs.

It has been proposed that the presence in human breast tumors of oestrogen receptor and milk proteins, especially α-lactalbumin, may indicate an intact oestrogen-receptor mechanism. Therefore synthesis of α-lactalbumin has been proposed to be a marker for tumors responsive to hormone therapy [16–18]. Because of inconsistencies in the immunoassay procedure, reports about the presence of milk proteins in tumor tissue were inconclusive [19]. In the following it was shown by cDNA hybridization that α-lactalbumin mRNA was not present in human mammary tumors [20]. In general cloned cDNA probes would be superior to immunological methods for the detection of specific markers in tumor tissue and tumor cell lines. The availability of a complete set of cDNA probes for all abundant lactation-specific mRNAs opens up the possibility of finding markers for steroid-hormone-responsive tumors. Since the mouse offers excellent genetic and experimental versatility for the study of mammary tumorology we characterized and cloned the lactation-specific mRNAs of the mouse.

Abbreviations. AMV, avian myeloblastosis virus; cDNA: DNA complementary to mRNA; NaCl/Cit, 150 mM sodium chloride, 15 mM sodium citrate, pH 7.0; Hepes, 4-(2-hydroxyethyl)-1-piperazinecthane sulfonic acid.

Enzymes. Reverse transcriptase or RNA-directed deoxyribonucleotidyltransferase (EC 2.7.7.49); DNA polymerase (Kornberg enzyme) (EC 2.7.7.7); nuclease S1 (EC 3.1.30.1); terminal deoxynucleotidyltransferase (EC 2.7.7.31); ribonuclease (EC 3.1.27.5).

MATERIALS AND METHODS

Materials

The radiochemicals and the rabbit reticulocyte translation system *in vitro* were obtained from Amersham, nitrocellulose filters were from Schleicher & Schüll, guanidine-HCl and guanidine-SCN were from Fluka, sucrose, heparin, sarcosyl and Triton X-100 were from Serva, poly-

nucleotide kinase and terminal deoxynucleotidyltransferase were from PL Biochemicals and oligo(dT)-cellulose type 2 and synthetic *Hind*III linker molecules were from Collaborative Research. Restriction endonucleases, *Escherichia coli* DNA polymerase I (Kornberg enzyme), *Escherichia coli* DNA polymerase I (large fragment), T4 DNA ligase, deoxymononucleoside triphosphates, *E. coli* tRNA and RNase A were obtained from Boehringer Mannheim; salmon sperm DNA and *Aspergillus oryzae* S1 nuclease, were bought from Sigma. Methylmercury(II) hydroxide was obtained from Alfa, oxytocin (Orasthin) was from Hoechst AG and *Staphylococcus aureus* protein A (Pansorbin) was obtained from Calbiochem-Behring Corp. Avian myeloblastosis virus reverse transcriptase was provided by J. Beard through the Office of Program Resources and Logistics, National Cancer Institute of the USA.

Isolation of Mouse and Rat Milk; Preparation of Antibodies against Mouse α-Casein and Mouse β-Casein

Lactating animals, 5–15 days after parturition, were separated from their litter for 5 h, 15 min after intraperitoneal injection of oxytocin, 1 I.U. for the mouse and 10 I.U. for the rat, the animals were milked by hand. The milk was stored until use at −20 C. Skimmed milk was prepared by centrifugation and removal of the cream accumulating on top of the milk. The caseins were isolated from skimmed mouse milk [21] and separated according to size on sodium dodecylsulfate/polyacrylamide gels [22]. For immunization of rabbits, α1-casein (43 kDa) and β-casein (26 kDa) were eluted from gel slices with a solution containing 300 mM NaCl and 0.1% sodium dodecylsulfate. Before bleeding rabbits to death the immune response was monitored with Ouchterlony double-diffusion tests [23]. Human milk was kindly provided by C. Lohmann.

Isolation of mRNA from Lactating Mammary Glands

Mammary gland tissue was dissected from animals killed by cervical dislocation. After quick freezing in liquid nitrogen, the tissue was stored at −70 C. Total polysomal RNA was isolated as described previously [24] with the exceptions that the preparation of polysomes was in the absence of sodium deoxycholate and in the presence of 100 μg cycloheximide/ml polysome buffer. For the isolation of total cellular RNA the guanidine thiocyanate method of Chirgwin et al. [25] was applied. Poly(A)-containing RNA was enriched by oligo(dT)-cellulose chromatography [26].

Synthesis of Double-Stranded cDNA, Construction of Recombinant Plasmids, Transformation and Screening for Recombinant Plasmids

Double-stranded cDNA was synthesized as described earlier [27] with modifications. The first DNA strand was synthesized in a 100-μl assay containing 50 mM Tris/HCl, pH 8.3, 140 mM KCl, 10 mM $MgCl_2$, 30 mM 2-mercaptoethanol, 10 μg $(dT)_{12-18}$, 0.5 mM each of dATP, dGTP, dTTP, [α-^{32}P]dCTP (1.1 Ci/mmol), 16 U AMV reverse transcriptase and 10 μg mRNA. Second-strand synthesis was directed either by AMV reverse transcriptase [27] or by *E. coli* DNA polymerase I [28]. S1-digested double-stranded cDNA hairpin molecules were inserted either via *Hind*III linkers essentially as described earlier [27] or via dCMP homopolymer tailing into linearized dGMP-homopolymer-tailed [29] pUR51 vector DNA [30]. After ligation or annealing, the recombinant plasmids were used for transformation of *E. coli* strain χ1776 /AM15 [30, 31]. Recombinants of pUR51 were selected for ampicillin resistance and white colonies on indicator plates [30]. Clones carrying recombinant plasmids, which had integrated a double-stranded cDNA corresponding to abundant, lactation-specific mRNAs, were detected by colony hybridization according to Grunstein and Hogness [32]. As hybridization probe for prescreening either radioactively labelled single-stranded cDNA prepared from total mammary gland mRNA or radioactively labelled single-stranded cDNA prepared from size- fractionated mammary gland mRNA was used.

Isolation of cDNA-clone-selected mRNA; Translation and Immunoprecipitation in vitro

Plasmid DNA was isolated either from a 1-ml overnight culture or from a single colony using the alkaline extraction procedure of Birnboim and Doly [33]. Plasmid DNA was bound to nitrocellulose filters [34] and mRNA complementary to the cloned cDNA was isolated from total cellular RNA [35]. Total mRNA and cDNA-clone-selected mRNA were translated either in the rabbit reticulocyte lysate system, applying conditions specified by the supplier, or in the wheat germ cell-free translation system prepared according to Roberts and Patterson [36]. Cotranslational processing of the *in vitro* translocation products was achieved by adding 8 A_{280} units dog pancreatic microsomal membranes to 1 ml of the translation assay [37–39]. Immunoprecipitation of *in vitro* translation products was according to Lingappa et al. [39].

Gel Electrophoresis Elution of RNA

Proteins were separated on 15% or 13–20% sodium dodecylsulfate/polyacrylamide gels essentially as described [22]. Fluorography, using PPO in dimethylsulfoxide, was done according to Laskey and Mills [40]. Messenger RNA was size fractionated on urea/citrate/agarose gels [41] and eluted in a high-salt buffer [25]. For size determination of mRNA 8 mM methylmercury(II) hydroxide/agarose gels were used [42]. Electrophoresis of restricted or unrestricted plasmid DNA was carried out either in agarose in Tris/phosphate buffer [43] or in polyacrylamide gels in Tris/borate buffer [44]. DNA and RNA were visualized by staining with ethidium bromide.

RNA Blotting and Hybridization with Cloned cDNA Probes

RNA was blotted from methylmercury(II) hydroxide/agarose gels to nitrocellulose filters and hybridized [45] to nick-translated plasmid DNA [46].

RESULTS

Mouse and Rat Milk Proteins

The description of major milk proteins has been confined to a few species, notably to the cow and the ewe [47] as well as the guinea-pig [48, 49].

In order to illustrate the protein complexity in mouse milk and in mouse mammary gland tissue, the proteins of the respective sources were separated on sodium dodecylsulfate polyacrylamide gels and stained with Coomassie blue. The

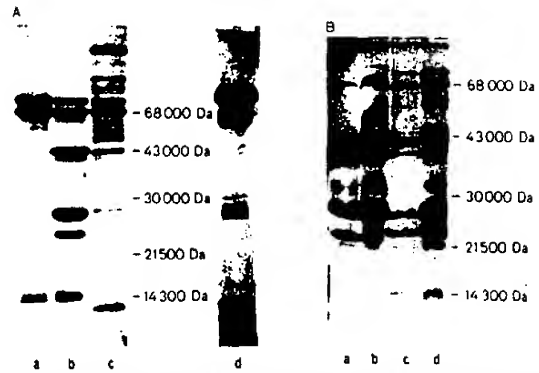


Fig. 1. *Sodium dodecylsulfate polyacrylamide gel pattern of mouse, rat and human milk proteins.* Proteins were precipitated by adding one volume 10% trichloroacetic acid. After 10 min on ice the proteins were pelleted by centrifugation for 5 min in an Eppendorf centrifuge, washed with acetone, dried under reduced pressure, boiled for 3 min in 10 mM 2-mercaptoethanol-containing sample buffer and separated on a 13–20% sodium dodecylsulfate polyacrylamide gel (A) or on a 15% sodium dodecylsulfate polyacrylamide gel (B). (A) (a) Mouse whey proteins, (b) total mouse milk proteins, (c) proteins from mouse breast homogenate, (d) total human milk proteins; (B) (a) 40 μg total rat milk protein, (b) 100 μg total rat milk protein, (c) 30 μg total mouse milk protein, (d) 75 μg total mouse milk protein. Lysozyme (14.3 kDa), soybean trypsin inhibitor (21 kDa), carbonic anhydrase (30 kDa), ovalbumin (43 kDa) and bovine serum albumin (68 kDa) was used as size markers

pattern of mouse milk shows seven more abundant proteins and one less abundant protein (Fig. 1 A, lane b and Fig. 1 B, lane d). Conventionally the milk proteins are classified into the acid-precipitable caseins and the whey proteins [21]; the latter consisting of α-lactalbumin, milk serum albumin and lactoferrin, and in the case of the whey of ruminants, also of β-lactoglobulin [21, 50].

Acidification of mouse milk to pH 4.5 with 1 M HCl leads to the precipitation of caseins, while the whey proteins remain in solution [50]. The mouse caseins have, on sodium dodecylsulfate/polyacrylamide gels, an apparent molecular mass of 43 kDa, 39 kDa, 26 kDa, 23.7 kDa and 21 kDa (Fig. 1 A, lane b and Fig. 1 B, lane d) and the whey proteins have an apparent molecular mass of 78 kDa, 67 kDa, and 14.3 kDa (Fig. 1 A, lane a). Because of a later-described immunological cross-reactivity as well as a later-found cross-hybridization of their mRNAs the 43-kDa and 39-kDa proteins were termed α1 and α2-casein. The other three caseins were named β, γ, and δ-casein with decreasing molecular weight. The whey proteins have been characterized fairly well in cow [50], sheep [50] and guinea-pig [49]. In those species α-lactalbumin with a molecular mass of 14.3 kDa appears to be a major whey protein. Mouse whey consists of a few different protein species, as can be judged from the sodium dodecylsulfate/polyacrylamide gel pattern (Fig. 1 A, lane a). Based on an enzymatic assay, specific for α-lactalbumin, the 14.3-kDa protein had been assigned to α-lactalbumin [51]. Recently it was shown that a novel whey protein exists which is about twice as abundant as α-lactalbumin, which comigrates with α-lactalbumin on sodium dodecylsulfate/polyacrylamide gels [52] and which was named whey acidic protein [53]. By comparison with the whey proteins from other species the 67-kDa and 78-kDa protein could very well be milk serum albumin [21] and lactoferrin [50] of the mouse. All eight abundant milk proteins can be correlated according to size to abundant proteins in the mouse mammary gland tissue homogenate. Beside these are several abundant proteins in the homogenate which do not have an equivalent in the milk. Predominant among those is a protein with the apparent size of about 13.7 kDa and a very large one with a molecular mass of more than 100 kDa (Fig. 1 A, lane c). In order to simplify the identification of the various mouse milk proteins their pattern was compared to the protein pattern in rat and human milk (Fig. 1 A, lane d, Fig. 1 B, lanes a and b). Rat milk proteins show a similar but not identical pattern to mouse milk: Apparently there are no proteins in rat corresponding in size to mouse α1-casein and the whey proteins of 14.3 kDa. The human milk protein pattern gives a good example for species diversification of apparent size and abundance of homologous proteins (Fig. 1 A, lane d, Fig. 1 B, lanes a and b).

Isolation of Mouse Mammary Gland mRNA

For translation *in vitro* and preparation of cDNA, mouse mammary gland cellular RNA and total polysomal RNA was prepared and enriched for poly(A)-containing mRNA by oligo(dT)-cellulose affinity chromatography. The isolated RNA fractions were applied to partial denaturing urea/citrate/agarose gels and visualized by ethidium bromide staining in order to estimate the integrity of the different types of RNA molecules (Fig. 2). From the intensities of 28-S and 18-S rRNA in the total mouse mammary RNA (Fig. 2, lane d), in its poly(A)-lacking RNA fraction (Fig. 2, lane f) and in total polysomal RNA (Fig. 2, lane g) it was possible to estimate the degree of degradation during the various isolation procedures. As a result we found that very little degradation occurred during direct isolation of cellular RNA from gland tissue using guanidinium thiocyanate as denaturing agent [25]. We did not succeed in getting the same degree of undegraded RNA when polysomes were used as intermediates for the isolation of mRNA. This might be due to the reported high content of RNase activity in milk [21]. The size distribution of poly(A)-plus RNA from total cellular RNA revealed the presence of a few abundant mRNAs (Fig. 2, lane e) with the

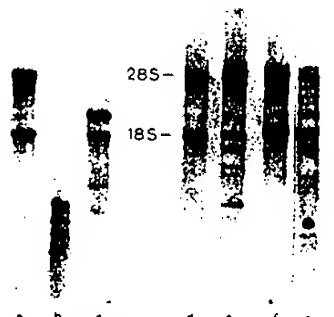


Fig. 2. *Urea citrate agarose gel of mouse mammary gland mRNA.* 15 µg RNA were loaded in one well (2 × 8 mm). After electrophoresis the RNA was stained with ethidium bromide (0.5 µg/ml) and visualized under ultraviolet light. (a) Rat liver total polysomal RNA, (b) chicken globin mRNA, (c) chicken oviduct mRNA, (d) mouse mammary gland total cellular RNA, (e) mouse mammary gland mRNA, (f) oligo(dT)-cellulose flow-through of mouse mammary gland total cellular RNA, (g) mouse mammary gland total polysomal RNA

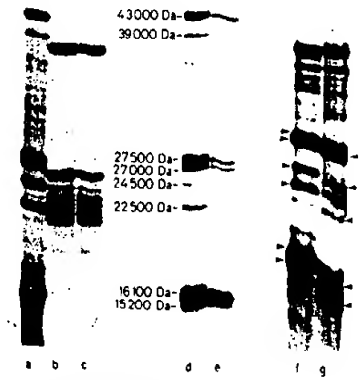


Fig. 3. *Sodium dodecylsulfate/polyacrylamide gel pattern of translation products* in vitro *of mouse and rat mammary gland mRNAs.* The mRNA was translated either in the rabbit reticulocyte lysate system or in the wheat germ system, supplemented with L-[^{35}S]methionine. Cotranslational processing was performed by the addition of 8 A_{280} units dog pancreatic microsomal membranes/ml translation assay. The newly synthesized proteins were separated on a 15%, sodium dodecylsulfate/polyacrylamide gel and visualized by autoradiography. (a) Mouse mammary gland total cellular RNA translated in the wheat germ system, (b) rat mammary gland mRNA translated in the rabbit system, (c) rat mammary gland total cellular RNA translated in the rabbit system, (d) mouse mammary gland total cellular translated in the rabbit system, (e) mouse mammary gland mRNA translated in the rabbit system, (f) mouse mammary gland mRNA translated in the wheat germ system in the absence of microsomes, (g) mouse mammary gland mRNA translated in the wheat germ system in the presence of microsomes. The arrows indicate the preproteins synthesized *in vitro* and the corresponding cotranslationally processed proteins

apparent size of 680 nucleotides, 1030 nucleotides, 1300 nucleotides and 1760 nucleotides in length. These distinct RNAs can be identified as mRNAs by their respective translation products *in vitro* and can be correlated to cDNAs of distinct sizes after reverse transcription.

Translation of Mouse and Rat Mammary Gland mRNA in vitro

As a first step in the process to identify individual mouse mammary gland mRNAs we compared the translation products of total mRNA *in vitro* with the pattern of abundant proteins in milk and tissue homogenate. Fig. 3 shows the translation products *in vitro* of mouse and rat mammary gland mRNA separated on a sodium dodecylsulfate/polyacrylamide gel. Both mRNA populations code for a few predominant proteins. The prominent proteins synthesized *in vitro* from the mouse mRNA have an apparent molecular mass of 43 kDa, 39 kDa, 27.5 kDa, 27 kDa, 24.5 kDa, 22.5 kDa, 16.1 kDa, 15.2 kDa; those of the rat of 38.5 kDa, 26.8 kDa, 24.5 kDa, 23 kDa and 22.5 kDa. For unknown reasons on some gels the 23-kDa and 22.5-kDa rat proteins do not separate and show up as a broad band.

The three caseins of the guinea-pig [49] as well as bovine αs1-casein and bovine β-casein [54, 55] do not contain any cysteine residues. κ-Caseins, on the other hand, contain two cysteine residues, which are important for their function *in vitro* [56]. Translation of mouse mammary gland mRNA *in vitro* in the presence of [^{35}S]cysteine showed, that γ-casein contains few cysteine residues and α-casein contains no cysteine residues. Mouse β-casein, on the other hand, contains cysteines (data not shown) and is different from the respective guinea-pig and bovine proteins.

Proteins which are synthesized in the mammary gland tissue and later appear in the milk must undergo secretorial processes, normally accompanied by cotranslational processing of an amino-terminal signal peptide [37, 38]. Therefore, we looked to see whether the abundant proteins synthesized *in vitro* were processed when microsomal membranes were added to the synthesizing system *in vitro*. Fig. 3 shows the *in vitro* translation products in the absence and presence of dog pancreatic microsomal membranes. We found that all abundant mouse proteins are cotranslationally processed to about 80%, a degree consistant with findings by others [39]. According to their size, the processed *in vitro* translation products can be fairly well correlated to specific milk proteins. In the mouse the 43-kDa and 39-kDa proteins synthesized *in vitro* (Fig. 3, lanes f and g) can be assigned to α1 and α2-casein. The 27.5-kDa and 27-kDa proteins (Fig. 3, lanes f and g), which are termed β1 and β2-casein because of a later-shown immunological cross-reactivity as well as a later-found cross-hybridization of their mRNAs, are processed to a molecular mass of 26 kDa and comigrate with β-casein of the milk (Fig. 1 A, lane b). The mouse γ and δ-casein in the milk, which have a molecular mass of 23.7 kDa and 21 kDa, can be correlated to the processed *in vitro* proteins of 23 kDa and 21 kDa (Fig. 3, lanes f and g). The two smallest *in vitro* translation products of mouse mammary gland mRNA, with a size of 16.1 kDa and 15.2 kDa are processed to a molecular weight of 14.5 kDa and about 13.7 kDa respectively (Fig. 3, lanes f and g). In the milk, however, only one small protein of 14.3 kDa can be seen on sodium dodecylsulfate/polyacrylamide gels (Fig. 1 A, lane b). Applying ion-exchange chromatography as a further purification step of the 14.3-kDa milk protein, it had been shown recently that at least two proteins

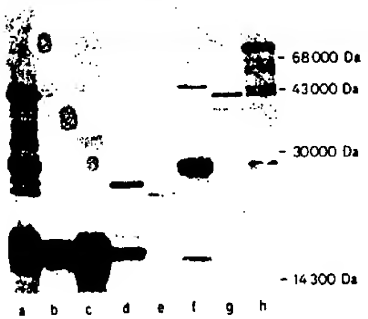


Fig. 4. *Sodium dodecylsulfate/polyacrylamide gel pattern of translation products* in vitro *of size-fractionated mouse mammary gland mRNA.* Size-fractionated mouse mammary gland mRNA was translated in the rabbit reticulocyte lysate system supplemented with L-[^{35}S]methionine. The newly synthesized proteins were separated on a 15%, sodium dodecylsulfate/polyacrylamide gel and visualized by autoradiography. Agarose-gel eluted mRNAs of the following size fractions were used for translation *in vitro*: (b) 560–640 nucleotides, (c) 650–700 nucleotides, (d) 800–1040 nucleotides, (e) 1100–1250 nucleotides, (f) 1300–1500 nucleotides, (g) 1550–1780 nucleotides, (h) 1800–2800 nucleotides. Lane (a) shows the translation products of total cellular mRNA *in vitro*. Lysozyme (14.3 kDa), carbonic anhydrase (30 kDa), ovalbumin (43 kDa) and bovine serum albumin (68 kDa) were used as size markers. The 36 kDa protein, appearing in each lane, is synthesized in the rabbit reticulocyte lysate system independently of the mRNA added

comigrate at this position on sodium dodecylsulfate/polyacrylamide gels. A minor one is α-lactalbumin [52] and a major one is an acidic whey protein [53], possibly identical to the 13.7-kDa protein product processed *in vitro*.

As we could show by sequence analysis (L. G. Hennighausen, A. Steudle, and A. E. Sippel, unpublished results) the 14.5-kDa protein synthesized *in vitro* is not α-lactalbumin but a small casein, which we named ε-casein. The four abundant translation products of rat mammary mRNA *in vitro* undergo cotranslational processing to sizes comparable to the abundant proteins in the rat milk (Fig. 1 B, lanes a and b).

Translation of Size-Fractionated mRNA in vitro

Translation of size-fractionated mRNA *in vitro* can be used to correlate individual mRNAs to their respective abundant mouse mammary gland proteins. For this purpose poly(A)-rich RNA was separated on urea/citrate/agarose gels. Taking into consideration the visible mRNA bands, gels were sliced, the RNA was eluted from gel slices and translated. Fig. 4 shows translation products of the mRNA fractions with increasing molecular weight (Fig. 4, lanes b–h).

When these products were compared to the *in vitro* translation products of total mRNA (Fig. 4, lane a) the following results could be derived. The RNA fraction of the size 560–1040 nucleotides contains the mRNAs for the two small proteins of 15.2 kDa (whey acidic protein) and 16.1 kDa (ε-casein) (Fig. 3, lane a). Because of broad protein bands appearing particulary when mRNA fractions below 700 nucleotides were translated no correlation of mRNA size and protein product size is possible in this region. The mRNA fraction from 800 to 1040 nucleotides long codes for ε-casein, γ-casein and a minor protein of 18.3 kDa. A comparison of the products synthesized *in vitro*, coded on mRNAs of the size 800–1040 nucleotides and 1040–1250 nucleotides, clearly shows that γ-casein, a protein of 24.5 kDa, is coded for by a smaller mRNA than δ-casein, a protein of 22.5 kDa. This deviation from the direct correlation of mRNA size and protein product size was later confirmed by cDNA cloning (see Fig. 9) and is evidence for large differences in the proportion of the lengths of coding and non-coding regions in the mRNAs for γ and δ-casein. The mRNA fraction of 1300–1500 nucleotides in size codes for the two β-caseins, the mRNA fraction of 1500–1780 nucleotides for the two α-caseins. The abundant translation product of the 1800–2800-nucleotide fraction is a protein of about 77 kDa possibly the iron-binding protein lactoferrin.

Messenger RNA size fractionation done in this way leads to a high degree of enrichment of the specific abundant mRNAs and to a useful separation from each other. It opens up the opportunity to use these mRNA fractions for the synthesis of specific cDNA hybridization probes, which can be used to prescreen a cDNA library in cloning experiments.

Immunoprecipitation of α and β-Casein Synthesized in vitro

Our finding that mouse mammary gland mRNA stimulated *in vitro* the production of possibly two α-caseins (α1, 43 kDa; α2, 39 kDa) and two β-caseins (β1, 27.5 kDa; β2, 27 kDa) initiated immunological experiments to prove the relationship between the protein products synthesized *in vitro* and the respective milk proteins. We isolated from milk the α-casein of 43 kDa and β-casein by elution of sodium dodecylsulfate/polyacrylamide gel slices. Antisera prepared against the 43-kDa α1-casein showed a double precipitation line with total milk proteins in Ouchterlony gel diffusion assays, giving a first indication for the presence of two antigens with cross-reacting determinants. When the same antisera were used to identify α-casein in the protein products of total mammary gland mRNA synthesized *in vitro* the 43 kDa together with the 39-kDa protein appeared in the immune precipitate (Fig. 5, lane a). The close relationship between these two proteins is independently reinforced by our later finding that cloned α-casein cDNA hybridizes equally well to both α-casein mRNAs. Antisera against milk β-casein (26 kDa) immunoprecipitated both the 27.5 kDa and the 27 kDa *in vitro* translation product (Fig. 5, lane c). Again, as we show later, the relationship of both proteins is indicated by cross-hybridization of their mRNA-specific cloned cDNA sequences.

Cloning of Mouse Mammary Gland mRNA Sequences

Mouse mammary gland poly(A)-rich RNA was reverse-transcribed into cDNA using AMV reverse transcriptase. Second-strand synthesis was done either with reverse transcriptase after alkali digestion of the mRNA or with *E. coli* DNA polymerase I after heat denaturation of the mRNA · cDNA hybrids. S1 treatment of the hairpin DNA preceeded the insertion into cloning vectors. The efficiency of the various DNA synthesis steps was controlled by incorporation of radioactive deoxynucleotide precursors and the degree of S1 resistance. The quality of the DNA intermediates was controlled by autoradiography after agarose gel electrophoresis. We found that 30% of the mass of mRNA was transcribed into cDNA, that roughly 60% of this cDNA became S1 resistant during second-strand synthesis, that about 30%

Table 1. *Data of mouse milk proteins, their respective mRNAs and cloned cDNAs*

Protein specificity	Occurrence milk	Apparent protein size			Apparent length		Name	Recombinant plasmids			
		in vivo	*in vitro*	*in vitro* + microsomes	mRNA	mRNA-cDNA		construction by linker	G-C tails	length of insert	number of positive clones[b]
		Da			bases	base pairs				base pairs	
α1-Casein	+	43000	43000	43000	1600	1460	pmαCa1	X		750	12
α2-Casein	+	39000	39000	39000	1600	1460					
β1-Casein	+	26500	27500	26000	1450	1140	pmβCa2	X		1150	18
β2-Casein	+	26500	27000	26000	1450	1140	pmβCa1	X		450	
γ-Casein	+	23700	24500	22500	860	840	pmγCa1	X		590	2
δ-Casein	+	21000	22500	21000	1150	950	pmδCa1		X	880	10
ε-Casein	?[a]	?	16100	14500	860	790	pmεCa1		X	850	6
Whey acidic protein	+	13700	15200	13700	680	620	pmWAP1		X	460	12
							pmWAP1		X	450	

[a] ε-Casein has so far neither been isolated from mouse milk, nor is the apparent molecular weight of the *in vivo* protein known.

[b] The plasmid DNA of 99 cDNA clones, which hybridize to cDNA prepared from total mRNA, was further hybridized to the cDNA inserts of the 6 cloned cDNA specificities. The number of positives for each cDNA specificity is indicated.

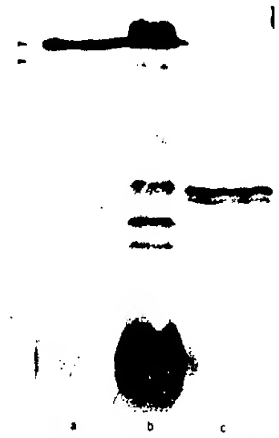


Fig. 5. *Immunoprecipitation of the mouse α and β-caseins synthesized* in vitro. Mouse α and β-caseins were immunoprecipitated from the *in vitro* translation products of mouse mammary gland total mRNA using antibodies prepared either against the 43-kDa α-casein or the 26-kDa β-casein of mouse milk. The immunoprecipitated proteins were separated on a 15% sodium dodecylsulfate/polyacrylamide gel as described in Materials and Methods. (a) Immunoprecipitated α-caseins; (b) *in vitro* translation products of mouse mammary gland total mRNA after immunoprecipitation of the β-caseins; (c) immunoprecipitated β-caseins. The arrows indicate the two immunoprecipitated α-caseins

stayed S1 resistant after heat denaturation and that about 35% of the original cDNA label could be finally recovered after digestion of the hairpin DNA with S1 nuclease.

After first-strand synthesis some abundant cDNA·mRNA hybrids of distinct size are detectable upon gel electrophoretic analysis (data not shown). Considering the mRNA sizes of the various mammary gland proteins, as determined on urea citrate/agarose gels and by blot analysis (see later) the different hybrids can be assigned to the mRNAs of the individual mammary gland proteins (Table 1). The apparent sizes of the double-stranded cDNA molecules are slightly smaller than those of the respective hybrids (Table 1), indicating, that the second-strand synthesis is not complete. This can be explained in view of the well known fact that the synthesis of the second strand is initiated by a fold-back loop at the 3′ end of the single-stranded cDNA. S1 nuclease digestion and oligodeoxycytidine tailing with terminal deoxynucleotidyltransferase leads to no further alteration of the apparent size of the double-stranded cDNA.

An interesting phenomenon can be observed concerning the cDNA for α-casein. Whereas the α-casein-specific mRNA·cDNA hybrid is of homogeneous size, the double-stranded cDNA splits into two populations, a major one around 1360 base pairs and a minor one around 1460 base pairs in length (data not shown). This is consistant with the possibility that α-casein cDNA consists of two species of equal length, which form basically two terminal hairpins of different size during second-strand synthesis.

We used two different methods for the ligation of double-stranded cDNA to vector DNA (see Materials and Methods). Our cloning efficiency in *E. coli* χ1776 in both cases was roughly 3×10^4 recombinants/μg double-stranded cDNA. Recombinants were detected in one step by a rapid combined antibiotic resistance and colour-indicator screening method [30].

Identification of Clones Corresponding to Specific mRNAs

A fraction of bacterial clones from the mouse mammary gland cDNA library was screened for their content of cDNA sequences corresponding to abundant lactation-specific mRNAs by three successive identification steps.

A total of 239 independent colonies were grown on nitrocellulose filters, lysed and hybridized [32] with radioactive cDNA prepared from total mammary gland mRNA. 69 colonies (29%) gave intense signals on the autoradiograph and 30 colonies (12.5%) gave weaker signals. In a second screening step we used radioactive cDNA prepared from size-fractionated mRNA templates. These hybridization probes were highly enriched and rather specific for individual abundant mRNAs, which we had shown by translation *in vitro* (Fig. 4). When the 99 prescreened colonies were grown on replicaplated nitrocellulose filters and hybridized to these cDNA probes they could be divided in 7 groups, each enriched for clones corresponding to one or two abundant mRNA species. A more definite assignment of the cDNA clones to specific mRNAs was made by translation *in vitro* of mRNA complementary to plasmid DNA of individual colonies.

Messenger RNA retained by filter-bound plasmid pαCa1 directs the synthesis of both the 43-kDa and 39-kDa α-casein *in vitro* (data not shown). The mass ratio of the two α-caseins synthesized *de novo* was identical to the mass ratio derived by translation of total mRNA (Fig. 7A). This demonstrates that, under the conditions used, both α-casein mRNAs hybridize equally well to pαCa1. The incomplete DNA inserts of 8 out of a total of 12 α-casein clones were mapped with restriction enzymes and showed homology as if they were derived from the same mRNA possibly the mRNA for the at least ten times more abundant 43-kDa α-casein.

Cross-hybridization of pαCa1 to both mRNAs is further evidence for the already detected immunological relationship of the two α-caseins. Plasmid DNA of all β-casein-specific cDNA clones filter-bind mRNAs which code for a 27.5-kDa and a 27-kDa protein (Fig. 6, lanes d and e). After addition of a microsomal membrane fraction to the translation system *in vitro*, both proteins comigrate at a molecular mass of 26 kDa (Fig. 3). When the 27.5-kDa and 27-kDa translation products of total mRNA or RNA selected by β-casein cDNA clones were compared, the ratio between them remained constant (Fig. 7B). When we tested all eight β-casein-specific plasmid DNA inserts for their restriction enzyme cleavage pattern and the ratio with which they retained both β-casein mRNAs we found that seven were from the same mRNA. One of them, pβCa1, had a non-overlapping restriction map and selected a differt ratio of the two β-casein mRNAs (Fig. 6, lanes d and e and Fig. 7B).

The relative distribution of cDNA specificities in the mammary-gland-specific cDNA library was determined by a rescreen using radioactive nick-translated insert DNA from selected recombinant plasmids. The number of clones determined in this way are shown in Table 1 and can be a rough estimate of individual mRNA concentrations.

Characterization of cDNA Clones

Cloned cDNAs are a most helpful tool for the elucidation of the structure of mRNAs. They allow one, upon sequence analysis, to determine protein-coding and non-coding regions and to deduce the amino acid sequence of their respective protein specificities. The knowledge of their sequence is necessary to work out the exact exon/intron pattern of their respective genes in genomic DNA.

In order to prepare for sequence analysis, the plasmid DNAs of the cDNA clones were characterized by determina-

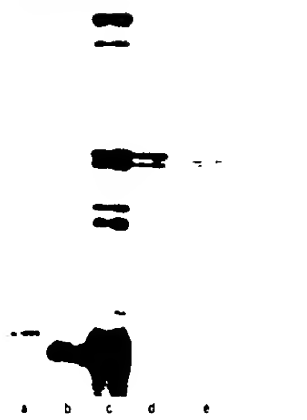


Fig. 6. *Sodium dodecylsulfate polyacrylamide gel pattern of* in vitro *translation products of clone-selected mouse mammary gland mRNAs.* Recombinant plasmids, immobilized on nitrocellulose filters, were used to select complementary mRNAs. Translation products in the rabbit reticulocyte lysate system of mRNAs complementary to following recombinant plasmids were separated on a 15% sodium dodecylsulfate polyacrylamide gel: lane (a) pαCa1, (b) pWAP1, (d) pβCa1, (e) pβCa2; in lane (c) the translation products of mouse mammary gland total mRNA *in vitro* are shown

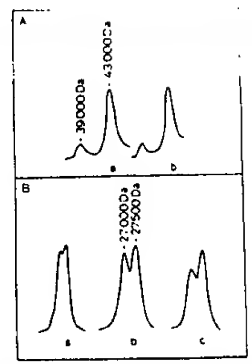


Fig. 7. *Relative concentrations of the two mouse α-caseins synthesized* in vitro *and the two mouse β-caseins synthesized* in vitro. [^{35}S]Methionine-labelled *in vitro* translation products, either of mouse mammary gland total mRNA or cDNA-clone-selected mRNA, were separated according to size on 15% sodium dodecylsulfate polyacrylamide gels and visualized by autoradiograph (Fig. 6). The autoradiograph was scanned to determine the relative concentrations of the two α-caseins synthesized *de novo* and the two β-caseins synthesized *de novo*. The following scanning profiles are shown: (A) (a) α1 and α2-casein (Fig. 6, lane c); (b) α1 and α2-casein, synthesized from clone-selected mRNA (autoradiograph not shown); (B) (a) β1 and β2-casein, synthesized from clone-selected mRNA (Fig. 6, lane e); (b) β1 and β2-casein, synthesized from total mammary gland mRNA (Fig. 6, lane c); (c) β1- and β2-casein, synthesized from clone-selected mRNA (Fig. 6, lane d)

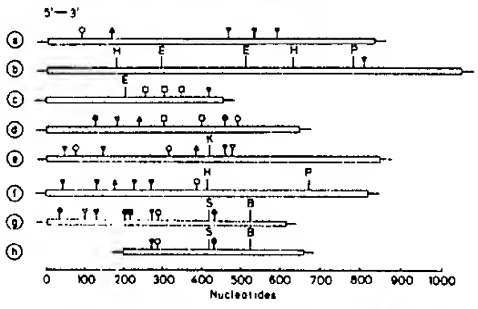


Fig. 8. *Restriction-endonuclease-cleavage maps from the mouse-mammary-gland-specific cDNAs.* Inserts of recombinant plasmids covering the longest overlapping stretch of the corresponding mRNA are shown. The 5'→3' orientation of the double-stranded cDNA insert within the vector was determined by sequence analysis as described in Results. (a) pmαCa1, (b) pmβCa2, (c) pmβCa1, (d) pmγCa1, (e) pmδCa1, (f) pmαCa1, (g) pmWAP1, (h) pmWAP2. The following symbols were used: (B) *BamH*I, (K) *Kpn*I, (▽) *Dde*I, (E) *Eco*RI, (○) *Sau*3A, (▼) *Hin*fI, (H) *Hin*dIII, (●) *Hpa*II, (□) *Taq*I, (P) *Pst*I, (△) *Alu*I, (■) *Ava*II, (S) *Sal*I, (▲) *Hae*III, (◇) *Hin*cII

tion of length, orientation and restriction enzyme cleavage map of their DNA insert. The 5'→3' orientation of the mRNA sequence was determined by localization of the mRNA-specific poly(A) tail or the longest open-reading frame via DNA sequencing (unpublished results). Restriction enzyme mapping helped us to align overlapping incomplete cDNA sequences. Table 1 gives data of those of our clones which cover together the longest overlapping stretch of mRNA sequence. Fig. 8 shows restriction maps of the milk-protein-specific clones, which are mentioned in Table 1.

Determination of mRNA Sizes Using Cloned cDNA Hybridization Probes

Cloned cDNAs were in all cases not full-length copies of their respective mRNAs. True mRNA sizes, of those specificities which are present in cloned cDNA, were therefore measured by a filter-hybridization procedure [45]. Total mammary gland mRNA of the mouse and rat was separated according to size on denaturing methylmercury(II) hydroxide/agarose gels, transferred onto nitrocellulose filters and hybridized with nick-translated radioactive DNA of individual recombinant plasmids. Fig. 9 shows RNA blots hybridized to the six cloned cDNA specificities. The DNA of α-casein cDNA clone pmαCa1 hybridizes to a 1600-nucleotide mouse mRNA and to a rat mRNA of equal or slightly smaller size. Within the limits of the resolution power of this technique mouse α1 and α2-casein mRNAs show similar, if not identical, length and β-casein mRNAs of mouse and rat show an equal length of 1450 nucleotides. Since β-caseins synthesized *in vitro* have a molecular mass of about 27 kDa, corresponding to a coding region of approximately 800 nucleotides, this result points to a considerably long non-coding region in the β-casein mRNAs. Mouse δ-casein mRNA must be another mRNA species with a substantial non-coding region. Already on the partial denaturing urea/citrate/agarose gels we had noticed a reverse correlation between mRNA length and protein size for γ and δ-casein. The apparent size of the mRNA for the 22-kDa δ-casein was larger than that for the 24.5-kDa γ-casein. The size determination on total denaturing methylmercury(II) hydroxide gels confirmed this observation. δ-Casein is coded for by a mRNA of 1150 nucleotides and γ-casein by a mRNA of 860 nucleotides. The same size of 860 nucleotides was also measured for rat γ-casein and for mouse ε-casein. Whey acidic protein, the smallest of the abundant mammary gland proteins, is coded on a mRNA with the size of 620 nucleotides in mouse and rat. Determination of the cDNA sequences proved that pαCa1 and pWAP1 contained the entire coding region for their corresponding proteins [58] (L. G. Hennighausen, A. Steudle, and A. E. Sippel, unpublished results).

Fig. 9. *Size determination of the lactation-specific mouse mammary gland mRNAs and their equivalents in rat.* 15 μg total cellular RNA from mouse and rat mammary gland were separated on methylmercury(II) hydroxide/agarose gels. The RNA was blotted onto nitrocellulose filters, hybridized to ^{32}P-labelled cloned cDNA probes and visualized by autoradiography. Hybridization was performed with pαCa1 and pWAP1 to mouse mRNA (a), with pβCa2 to mouse mRNA (b), with pγCa1 to rat (c) and mouse (d) mRNA, with pεCa1 to rat (f) and mouse (g) mRNA and with pδCa1 to mouse mRNA (j). Lanes e and h show *Hin*fI-restricted pBR322

DISCUSSION

Several levels of selection were used to identify milk-protein-specific cDNA clones from a mammary gland cDNA library of lactating mice.

a) The cDNA library was prescreened with a cDNA probe corresponding to the total mammary gland mRNA. Clones giving positive signals were considered specific for abundant mRNAs and were selected for further screening.

b) For a second screening we used cDNA probes complementary to size-fractionated mRNAs. These mRNAs had been shown in cell-free translation assays to code for proteins which could be correlated to abundant proteins synthesized by total mammary gland mRNA. One indication for a peptide synthesized *in vitro* to be specific for a milk protein would be its processing *in vitro* in the presence of microsomes [37, 38]. Most products processed *in vitro* could be reasonably well size-correlated to *in vivo* milk proteins. In the case of α and β-caseins the mRNA assignment was done by immunoprecipitation of *in vitro* products with monospecific antibodies.

c) A more definitive assignment was achieved by analysis and comparison of *in vitro* translation products of clone-selected mRNA species with the *in vitro* translation products of total mRNA. By comparing the size and the relative abundance we could show that the cDNA clones correspond to abundant and lactation-specific mRNAs. The absolute and relative number of specific cDNA clones isolated from the cDNA library is further good evidence that major mRNAs have been cloned.

d) Recently the sizes of the lactation-specific mRNAs in the rat mammary gland have been determined [13]. Cross hybridization of our clones with rat mammary gland mRNAs of the same sizes is independent evidence that we have isolated clones corresponding to the milk proteins of the mouse.

Absolute criteria for a certain specificity can only come from sequence analysis. Amino acid sequences of mouse caseins are not yet determined. A classification of the mouse milk proteins would be most meaningful by comparison with the defined milk proteins from other mammalian species. Sequence analysis of clones from all specificities described are under way in our laboratory.

Caseins, the most abundant proteins in the milk, have been isolated and characterized from several species including guinea-pig [49], cow [50] and sheep [50]. Equivalent caseins of those species have a similar molecular weight and as a general feature they contain no cysteine residues. In the mouse the low cysteine content holds true only for α-casein. Besides differences in cystein content, the 43-kDa mouse γ-casein has a higher molecular mass than any of the caseins described so far. Immunoprecipitation of *in vitro* translation products of total mouse mammary gland mRNA with antibodies directed against α-casein and antibodies against β-casein, as well as *in vitro* translation products of cDNA-clone-selected mRNA, gave evidence for the existence of two immunologically cross-reacting α-caseins and two immunologically cross-reacting β-caseins of different size. The α-casein cDNA clones, characterized so far, hybridize equally well to the two equally sized α-casein mRNAs. Experiments using cDNA clones from various parts of the mRNA as hybridization probe and sequence analysis will show the nature of the relationship between the 43-kDa and the 39-kDa α-casein. We also present evidence that the two β-caseins are coded for by two mRNAs of the same length. We are presently sequencing the two different β-casein cDNA clones to reveal their sites of homology. The rat in comparison seems to express either single α-casein and β-casein species or more than one specificity of the same size. Heterogeneity within individual clones of caseins seems to be no exception. Using two-dimensional gel systems a heterogeneity was found within the guinea-pig precasein A and precasein B [59] synthesized *in vitro*.

Mouse α, β and γ-casein cDNAs cross-hybridize with the respective mRNAs from the rat under stringent hybridization criteria, demonstrating reasonably good conservation among those two species. Translation of mouse and rat total mammary gland mRNA *in vitro* reveals an additional protein of 22.5 kDa in both species. For the rat this protein has been shown to be α-lactalbumin [15]. Based on sequence analysis (L. G. Hennighausen, A. Steudle, and A. E. Sippel, unpublished results) in the mouse the 22.5-kDa preprotein can be correlated to α-casein. The mouse δ-casein cDNA clone does not cross-hybridize to a rat mammary gland mRNA. In addition to the four species of casein cDNA clones, a further abundant cDNA was isolated. The plasmid DNA of this clone was complementary to a 860-nucleotide mRNA coding for a protein of 16.1 kDa, which we call ε-casein. Characteristic for mouse ε-casein cDNA is a *Pst*I restriction site in the 3' non-coding region at a similar position to that found in mouse β-casein cDNA. In the rat an equivalent cDNA clone has been detected. Its specificity was designated β'-casein by Gupta et al. [60] because its *Pst*I restriction site was in similar position to that in rat β-casein cDNA. By sequence analysis we could show that mouse ε-casein contains amino acid stretches homologous to bovine αs2-casein, but not to bovine β-casein (L. G. Hennighausen, A. Steudle, and A. E. Sippel, unpublished results).

The recombinant plasmids of 12% of the cDNA clones correspond to a mRNA coding for a novel mouse acidic whey protein. The 15.2-kDa protein synthesized *in vitro* was processed by cotranslational addition of microsomal membranes to a molecular mass of about 13.7 kDa. The protein sequence was deduced from the sequence of the cloned cDNA. This novel acidic whey protein shows a strong relationship to a recently described [61, 62] class of small cysteine-rich proteins, which include the hypothalamic peptide hormone carrier protein, neurophysin, and shows no relationship to α-lactalbumin of other species [58]. Recently from mouse and rat milk a whey protein has been isolated that comigrates with α-lactalbumin on sodium dodecylsulfate/polyacrylamide gels and which is about three times as abundant as α-lactalbumin [52, 53] (and K. E. Ebner, personal communication). The first 17 amino acids of the partially sequenced protein from the rat (K.E. Ebner, personal communication) show a reasonable homology to the first 17 amino acids of the processed mouse whey acidic protein as deduced from the sequence of our cDNA clone pWAP1 [58]. Amino acid composition strongly suggests that the isolated mouse milk protein [53] is identical to the acidic whey protein we describe. A cDNA clone from the rat has been isolated which originally was described to be α-lactalbumin specific [13]. Sequence analysis revealed that the cloned rat cDNA was the equivalent of the cDNA, specific for mouse acidic whey protein, cloned by us. Sequence comparison of the rat and of the mouse cDNA sequence demonstrated a considerable sequence homology in the coding region and an even greater homology in the 3' non-coding region (L. G. Hennighausen, A. E. Sippel, A. A. Hobbs, and J. M. Rosen, unpublished results).

In humans, guinea-pigs and ruminants α-lactalbumin has a molecular mass of 14.3 kDa and it is one of the major whey proteins [49, 50, 64]. 40% of the clones from a lactating guinea-pig mammary-gland-specific cDNA library were said to code

for α-lactalbumin [14]. The relative concentrations of α-casein, β-casein and α-lactalbumin in rat milk were estimated to be roughly 60:10:6 [65]. The cDNA corresponding to rat α-lactalbumin has been cloned recently, showing that about 5‰ of the mRNAs specific for the lactating rat mammary gland code for α-lactalbumin [15]. In the mouse milk, however, α-lactalbumin seems to be a much less abundant protein comprising not more than 0.85‰ of the total milk proteins [52]. We are in the process of screening further cDNA clones from the tissue-specific cDNA library in order to isolate cDNA clones coding for this minor mouse milk protein.

α-Lactalbumin has previously been identified in the cytosol fraction of mammary tumors, using radioimmunoassay [16, 18] and by immunoprecipitation of the translation products of total mRNA *in vitro* from mammary gland tumor tissue [17]. Using, however, cloned α-lactalbumin cDNA as a hybridization probe it was demonstrated that α-lactalbumin transcripts are not detectable in human tissue [20]. We have shown that, in the mammary gland of at least the mouse and the rat, there are besides α-lactalbumin two additional and abundant proteins of almost the same size, κ-casein and whey acidic protein. It is therefore possible that the small milk-specific proteins, used for the preparation of antibodies, were not pure antigens. Experiments to verify whether the acidic whey protein is the unknown α-lactalbumin-sized protein, expressed in some human mammary tumors, are underway in our laboratory. The difficulty of discriminating between the expression of the various proteins can best be avoided by the future use of cloned, defined cDNA probes for the detection of hormone-regulated mRNA species in normal and neoplastic tissue. Furthermore the cloned cDNAs will enable us to isolated the respective genes and to study their regulation by steroid hormones and peptide hormones on all levels of nucleic acid metabolism.

We thank U. Rüther for providing the cloning vector pUR51, R. Craig for a human α-lactalbumin cDNA clone, J. Beard for the AMV reverse transcriptase, W. Stoffel for the dog pancreatic microsomal membranes, J. Schröder for the wheat germ translation system, A. Steudle for technical assistance, G. Küster for keeping the mice, K. Ebner and J. Rosen for communication of data before publication and I. Götz-Kriehl for typing the manuscript. This work was supported by a grant of the *Deutsche Forschungsgemeinschaft* (SFB 74) to A. E. S. and by a fellowship of the *Studienstiftung des Deutschen Volkes* to L. G. H. This manuscript was submitted in partial fulfillment of the graduate requirements for L. G. H. Cloning experiments were carried out under L2/B2 conditions as specified by *Richtlinien zum Schutz vor Gefahren durch in vitro rekombinierte Nukleinsäuren* of the BMFT of the Federal Republic of Germany. Casein cDNA clones described in this paper are free for distribution on request.

REFERENCES

1. Schütz, G., Nguyen-Huu, M. C., Giesecke, K., Hynes, N. E., Groner, B., Wurtz, T. & Sippel, A. E. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 617–724.
2. Bannerjee, M. R. (1976) *Int. Rev. Cytol. 47*, 1–97.
3. Topper, Y. J. & Oka, T. (1974) in *Lactation* (Larson, B. L. & Smith, V. R., eds) vol. 1, pp. 327–348, Academic Press, New York & London.
4. Matusik, R. J. & Rosen, J. M. (1978) *J. Biol. Chem. 253*, 2343–2347.
5. Houdebine, L.-M. & Gaye, P. (1975) *Mol. Cell. Endocrinol. 3*, 37–55.
6. Schuster, R. C., Houdebine, L.-M. & Gaye, P. (1976) *Eur. J. Biochem. 71*, 193–199.
7. Rosen, J. M., O'Neal, D. L., McHugh, J. E. & Comstock, J. P. (1978) *Biochemistry. 17*, 290–297.
8. Nakhashi, H. L. & Qasba, P. K. (1979) *J. Biol. Chem. 254*, 6016–6025.
9. Burdin, L. J., Parker, D., Craig, R. K., Getova, T. & Campbell, P. N. (1981) *Biochem. J. 194*, 999–1006.
10. Teysset, B. & Houdebine, L.-M. (1981) *Eur. J. Biochem. 114*, 597–608.
11. Teysset, B. & Houdebine, L.-M. (1981) *Eur. J. Biochem. 117*, 563–568.
12. Metha, N. M., Ganguly, N., Ganguly, R. & Bannerjee, M. R. (1980) *J. Biol. Chem. 255*, 4430–4434.
13. Richards, D. A., Rodgers, J. R., Supowitt, S. C. & Rosen, J. M. (1981) *J. Biol. Chem. 256*, 526–532.
14. Craig, R. K., Hall, L., Parker, D. & Campbell, P. N. (1981) *Biochem. J. 194*, 989–998.
15. Dandekar, A. M. & Qasba, P. K. (1981) *Proc. Natl Acad. Sci. USA 78*, 4853–4857.
16. Woods, K. L., Cove, D. H. & Howell, A. (1977) *Lancet*, 2, 14–16.
17. Hall, L., Craig, R. K. & Campbell, P. N. (1979) *Nature (Lond.) 277*, 54–56.
18. Woods, K. L., Cove, D. H., Morrison, J. M. & Heath, D. A. (1979) *Eur. J. Cancer, 15*, 47–51.
19. Stevens, U., Laurence, D. J. R. & Ormerod, M. G. (1978) *Clin. Chim. Acta. 87*, 149–157.
20. Hall, L., Craig, R. K., Davies, M. S., Ralphs, D. N. L. & Campbell, P. N. (1981) *Nature (Lond.) 290*, 602–604.
21. McKenzie, H. A. (1967) *Adv. Protein Chem. 22*, 56–221.
22. Laemmli, U. K. (1970) *Nature (Lond.) 227*, 680–685.
23. Ouchterlony, Ö. (1948) *Ark. Kemi Mineral. Geol. 26B*, 1.
24. Sippel, A. E., Hynes, N., Groner, B. & Schütz, G. (1977) *Eur. J. Biochem. 77*, 141–151.
25. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry. 18*, 5294–5299.
26. Aviv, H. & Leder, P. (1972) *Proc. Natl Acad. Sci. USA. 69*, 1408–1412.
27. Sippel, A. E., Land, H., Lindenmaier, W., Nguyen-Huu, M. C., Wurtz, T., Timmis, K. N., Giesecke, K. & Schütz, G. (1978) *Nucleic Acid Res. 5*, 3275–3294.
28. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem 253*, 2483–2495.
29. Roychoudhury, R., Jay, E. & Wu, R. (1976) *Nucleic Acid Res. 3*, 101–116.
30. Rüther, U. (1980) *Mol. Gen. Genet. 178*, 475–477.
31. Hanahan, D. & Meselson, M. (1980) *Gene, 10*, 63–67.
32. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl Acad. Sci. USA. 72*, 3961–3965.
33. Birnboim, H. C. & Doly, J. (1976) *Nucleic Acid Res. 76*, 1513–1523.
34. Kafatos, F. C., Jones, C. W. & Efstratiadis, A. (1979) *Nucleic Acid Res. 7*, 1541–1552.
35. Recchardi, R. P., Miller, J. S. & Roberts, B. E. (1979) *Proc. Natl Acad. Sci. USA. 70*, 2330–2334.
36. Roberts, B. E. & Patterson, B. M. (1973) *Proc. Natl Acad. Sci. USA. 70*, 2330–2334.
37. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol. 67*, 835–851.
38. Blobel, G. & Dobberstein, B. (1975) *J. Cell Biol. 67*, 852–862.
39. Lingappa, V. R., Lingappa, J. R., Psasas, R., Ebner, K. E. & Blobel, G. (1978) *Proc. Natl Acad. Sci. USA. 75*, 2338–2342.
40. Laskey, R. A. & Mills, A. D. (1975) *Eur. J. Biochem. 56*, 335–341.
41. Lehrbach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry. 16*, 4743–4751.
42. Bailey, J. M. & Davidson, N. (1976) *Anal. Biochem. 70*, 75–85.
43. Loening, U. E. (1969) *Biochem. J. 113*, 131–138.
44. Peacock, A. C. & Dingman, C. W. (1967) *Biochemistry*, 6, 1818–1827.
45. Thomas, P. (1980) *Proc. Natl Acad. Sci. USA. 77*, 5201–5205.
46. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977).
47. Jeness, R. (1974) in *Lactation* (Larson, B. L. & Smith, V. R., eds) vol. 3, pp. 3–107, Academic Press, London and New York.
48. Craig, R. K., Brown, P. A., Harrison, O. S., McIlreavy, D. & Campbell, P. N. (1976) *Biochem. J. 160*, 57–74.
49. Craig, R. K., McIlreavy, D. & Hall, R. L. (1978) *Biochem. J. 173*, 633–641.
50. McKenzie, H. A., ed. (1971) *Milk Proteins; Chemistry and Molecular Biology*, vol. 2, Academic Press, New York.
51. Nahamatsu, Y. & Oka, T. (1980) *Biochem. J. 185*, 227–237.

52. Zamierowski, M. M. & Ebner, K. E. (1980) *J. Immunol. Methods, 36*, 211–220.
53. Piletz, J. E., Heinlen, M. & Ganshow, R. E. (1981) *J. Biol. Chem. 256*, 11509–11516.
54. Mercier, J.-C., Grosclaude, F. & Ribadeau-Dumas, B. (1972) *Atlas of Protein Sequences and Structure* (Dayhoff, M. O., ed.).
55. Ribadeau-Dumas, B., Brignon, G., Grosclaude, F. & Mercier, J. C. (1972) *Eur. J. Biochem. 25*, 505–514.
56. Taboski, G. (1974) *Adv. Protein Chem. 28*, 1–210.
57. Reference deleted.
58. Hennighausen, L. G. & Sippel, A. E. (1982) *Nucleic Acids Res. 10*, 2677–2684.
59. Pascall, J. C., Boulton, A. P., Parker, D., Hall, L. & Craig, R. K. (1981) *Biochem. J. 196*, 567–574.
60. Gupta, P., Rosen, J. M., D'Eustachio, P. & Ruddle, F. H. (1982) *J. Cell Biol. 93*, 199–204.
61. Drenth, J., Low, B. W., Richardson, J. S. & Wright, C. S. (1980) *J. Biol. Chem. 255*, 2652–2655.
62. Drenth, J. (1981) *J. Biol. Chem. 256*, 2601–2602.
63. Reference deleted.
64. Findlay, J. B. C. & Brew, K. (1972) *Eur. J. Biochem. 27*, 65–86.
65. Qasba, P. M. & Nakhashi, H. L. (1978) *Proc. Natl Acad. Sci. USA. 75*, 4739–4743.

L. G. Hennighausen and A. E. Sippel, Institut für Genetik, Universität zu Köln, Weyertal 121, D-5000 Köln 41, Federal Republic of Germany

ATTACHMENT 2

United States Patent [19]

Meade et al.

[11] Patent Number: 4,873,316

[45] Date of Patent: Oct. 10, 1989

[54] ISOLATION OF EXOGENOUS RECOMBINANT PROTEINS FROM THE MILK OF TRANSGENIC MAMMALS

[75] Inventors: Harry Meade, Newton, Mass.; Nils Lonberg, New York, N.Y.

[73] Assignee: Biogen, Inc., Cambridge, Mass.

[21] Appl. No.: 65,994

[22] Filed: Jun. 23, 1987

[51] Int. Cl.[4] C07K 3/02; C07K 3/12; C07K 3/18; C12N 15/00

[52] U.S. Cl. 530/412; 530/360; 530/361; 530/833; 530/832; 530/416; 530/417; 530/418; 435/68; 435/172.1; 435/172.3; 435/240.2; 935/53; 935/55; 935/70; 935/111; 800/1; 536/27; 536/28; 536/29

[58] Field of Search 435/68, 172.1, 172.3, 435/226, 240.2; 530/832, 833, 412, 360, 361, 303; 800/1; 935/53, 55, 70; 536/27, 28, 29

[56] References Cited

U.S. PATENT DOCUMENTS

4,018,752	4/1977	Bühler et al.	530/382
4,229,342	10/1980	Mirabel	530/382
4,376,072	3/1983	Connolly	530/382
4,396,601	8/1983	Salser et al.	435/172.3
4,462,932	7/1984	Lonergan	530/382
4,644,056	2/1987	Kothe et al.	530/382
4,736,866	4/1988	Leder et al.	

FOREIGN PATENT DOCUMENTS

0117059	8/1984	European Pat. Off.	435/172.3
0247494	12/1987		
0264166	4/1988		
WO88/00239	1/1988	PCT Int'l Appl.	
WO88/01648	3/1988	PCT Int'l Appl.	

OTHER PUBLICATIONS

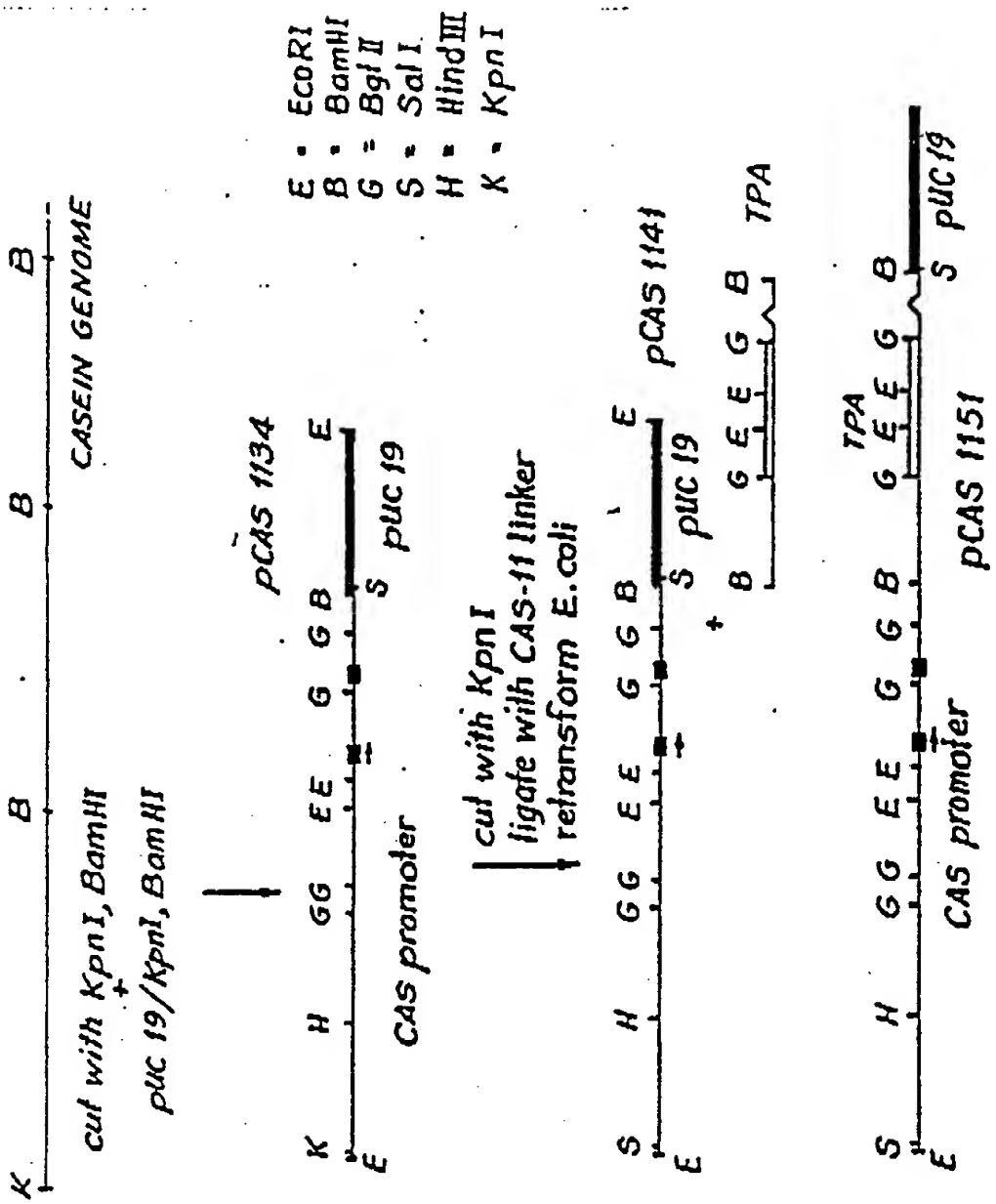
Gordon et al., Bio/Technology, 5, 1183-7, (Nov. 1987).
Lovell-Badge, Nature, 315, 628-629, (1985).
Hammer et al., Nature, 415, 680-683, (Jun. 20, 1985).
Garcia et al., Mol. Cell. Biol., 6(6), 1974-82, (1986).
Palmiter et al., J. Cell. Bioch., 8B, p. 25, Ab. #0890, (1984).
Fisher et al., J. B. C., 260 (20), 11223-11230, 1985.
Andres et al., Chem. Abs. 106(17):133024q, 1987 (for PNAS USA, 84(5), 1299-303, May 1987).
Ross et al., P.N.A.S. (USA), 82, 5880-84, 1985.
Brinster et al., Cell, 27, 223-31, Nov. 1981.

Primary Examiner—Margaret Moskowitz
Assistant Examiner—Jeff P. Kushan
Attorney, Agent, or Firm—James F. Haley, Jr.; Teresa L. Solomon

[57] ABSTRACT

This invention relates to the production of recombinant proteins in mammals' milk. Particularly, this invention relates to an expression system comprising the mammal's casein promoter which when transgenically incorporated into a mammal permits the female species of that mammal to produce the desired recombinant protein in or along with its milk. This invention also relates to the transgenic mammal that produces the desired recombinant product in its milk.

3 Claims, 1 Drawing Sheet


E = EcoR1
B = BamHI
G = Bgl II
S = Sal I
H = HindIII
K = Kpn I
CASEIN GENOME
cut with Kpn I, BamHI
pUC 19/Kpn I, BamHI
pCAS 1134
pUC 19
CAS promoter
cut with Kpn I
ligate with CAS-11 linker
retransform E.coli
pCAS 1141
TPA
pCAS 1151

ISOLATION OF EXOGENOUS RECOMBINANT PROTEINS FROM THE MILK OF TRANSGENIC MAMMALS

TECHNICAL FIELD OF INVENTION

This invention relates to the production of recombinant proteins in mammals' milk. Particularly, this invention relates to an expression system which comprises at least a milk-specific protein promoter operatively linked to a DNA sequence coding for a signal peptide and a desired recombinant protein product. When such a system is transgenically incorporated into a mammal, the recombinant protein is expressed in the milk of the lactating transgenic mammal. This invention also relates to the transgenic mammal that produces the desired recombinant product in its milk. Recombinant products produced by the expression systems and transgenically altered mammals of this invention can be produced at significantly less cost than by conventional recombinant protein production techniques.

BACKGROUND ART

Recombinant DNA technology has enabled the cloning and expression of genes encoding medically and agriculturally important proteins and glycoproteins. Such products include, for example, insulin, growth hormone, growth hormone releasing factor, somatostatin, tissue plasminogen activator, tumor necrosis factor, lipocortin, coagulation factors VIII and IX, the interferons, colony stimulating factor, the interleukins and urokinase.

Many of these important proteins, however, are large (molecular weights in excess of 30Kd), secreted, require sulfhydryl bonds to maintain proper folding, are glycosylated and are sensitive to proteases. As a result, the recombinant production of such products in prokaryotic cells has proven to be less than satisfactory because the desired recombinant proteins are incorrectly processed, lack proper glycosylation or are improperly folded. Accordingly, resort has been had to the production of those recombinant proteins in cultured eukaryotic cells. This technique has proven to be both expensive and often unreliable due the variability of cell culture methods. For example, average yields are 10 mg of recombinant protein per liter of culture media, with the resulting cost typically for exceeding $1,000 per gram of recombinant protein. Accordingly, resort has been had to the production of those recombinant proteins in cultured eukaryotic cells.

DISCLOSURE OF THE INVENTION

The present invention solves such problems by providing an efficient means of producing large quantities of recombinant protein products in the milk of transgenically altered mammals. According to this invention, a DNA sequence coding for a desired protein is operatively linked in an expression system to a milk-specific protein promoter, or any promoter sequence specifically activated in mammary tissue, through a DNA sequence coding for a signal peptide that permits secretion and maturation of the desired protein in the mammary tissue. More preferably, the expression system also includes a 3' untranslated region downstream of the DNA sequence coding for the desired recombinant protein. This untranslated region may stabilize the rDNA transcript of the expression system. Optionally, the expression system also includes a 5' untranslated region upstream of the DNA sequence coding for the signal peptide.

The expression system is transgenically introduced into a host genome by standard transgenic techniques. As a result, one or more copies of the construct or system becomes incorporated into the genome of the transgenic mammal. The presence of the expression system will permit the female species of the mammal to produce and to secrete the recombinant protein product, into or along with its milk. Such method permits the low cost, high level production of the desired proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 depicts the construction of a plasmid of this invention, pCAS 1151.

DEFINITIONS

As used in this application and claims, the terms recombinant protein and operatively linked have the following definitions:

Operatively linked—the linking of a milk-specific promoter or a promoter specifically activated in mammary tissue to a DNA sequence coding for a desired protein so as to permit and control expression of that DNA sequence and production of that protein.

Recombinant protein—a protein or peptide coded for by a DNA sequence which is not endogeneous to the native genome of the mammal in whose milk it is produced in accordance with this invention or a protein or peptide coded for by a DNA sequence which if endogeneous to the native genome of the mammal in whose milk it is produced does not lead to the production of that protein or peptide in its milk at the same level that the transgenic mammal of this invention produces that protein in its milk.

DETAILED DESCRIPTION OF THE INVENTION

This invention relates to processes, DNA sequences, compositions of matter and transgenic mammals for the production of recombinant proteins More specifically, this invention relates to the transgenic incorporation of one or more copies of a construct comprising a milk-specific protein promoter or any promoter sequence specifically activated in mammary tissue, operatively linked to a DNA sequence coding for a desired recombinant protein through a DNA sequence coding for a signal peptide that permits the secretion and maturation of the desired recombinant protein in the mammary tissue. The construct is transgenically incorporated into mammalian embryos and the recombinant protein product is subsequently expressed and secreted into or along with the milk of the lactating transgenic mammal.

Any mammal may be usefully employed in this invention. Preferably, mammals that produce large volumes of milk and have long lactating periods are preferred. Preferred mammals are cows, sheep, goats, mice, oxen, camels and pigs. Of course, each of these mammals may not be as effective as the others with respect to any given expression sequence of this invention. For example, a particular milk-specific promoter or signal sequence may be more effective in one mammal than in others However, one of skill in the art may easily make such choices by following the teachings of this invention.

Among the milk-specific protein promoters useful in the various embodiments of this invention are the casein promoters and the beta lactoglobulin promoter. The casein promoters may, for example, be selected from an alpha casein promoter, a beta casein promoter or a kappa casein promoter. Preferably, the casein promoter is of bovine origin and is an alpha S-1 casein promoter.

Among the promoters that are specifically activated in mammary tissue and are thus useful in accordance with this invention is the long terminal repeat (LTR) promoter of the mouse mammary tumor virus (MMTV) The milk-specific protein promoter or the promoters that are specifically activated in mammary tissue may be derived from either cDNA or genomic sequences. Preferably, they are genomic in origin.

Among the signal peptides that are useful in accordance with this invention are milk-specific signal peptides or other signal peptides useful in the secretion and maturation of eukaryotic and prokaryotic proteins. Preferably, the signal peptide is selected from milk-specific signal peptides or the signal peptide of the desired recombinant protein product, if any. Most preferably, the milk-specific signal peptide is related to the milk-specific promoter used in the expression system of this invention. The size of the signal peptide is not critical for this invention. All that is required is that the peptide be of a sufficient size to effect secretion and maturation of the desired recombinant protein in the mammary tissue where it is expressed.

Among the protein products which may be produced by the processes of this invention include, for example, coagulation factors VIII and IX, human or animal serum albumin, tissue plasminogen activator (TPA), urokinase, alpha-1 antitrypsin, animal growth hormones, Mullerian Inhibiting Substance (MIS), cell surface proteins, insulin, interferons, interleukins, milk lipases, antiviral proteins, peptide hormones, immunoglobulins, lipocortins and other recombinant protein products.

The desired recombinant protein may be produced as a fused protein containing amino acids in addition to those of the desired or native protein. For example, the desired recombinant protein of this invention may be produced as part of a larger recombinant protein in order to stabilize the desired protein or to make its purification from milk easier and faster The fusion is then broken and the desired protein isolated. The desired recombinant protein may alternatively be produced as a fragment or derivative of native protein or it may be produced having an amino acid sequence similar to the native protein. Each of these alternatives is readily produced by merely choosing the correct DNA sequence.

Preferably, the expression system or construct of this invention also includes a 3' untranslated region downstream of the DNA sequence coding for the desired recombinant protein. This region apparently stabilizes the RNA transcript of the expression system and thus increases the yield of desired protein from the expression system. Among the 3' untranslated regions useful in the constructs of this invention are sequences that provide a poly A signal. Such sequences may be derived, e.g., from the SV40 small t antigen, the casein 3' untranslated region or other 3' untranslated sequences well known in the art. Preferably, the 3' untranslated region is derived from a milk specific protein. The length of the 3' untranslated region is not critical but the stabilizing effect of its poly A transcript appears important in stabilizing the RNA of the expression sequence.

Optionally, the expression control sequences of this invention also include a 5' untranslated region between the promoter and the DNA sequence encoding the signal peptide. Such untranslated regions are preferably related to the promoter. However, they may be derived from other synthetic, semi-synthetic and natural sources. Again their specific length is not critical, however, they appear to be useful in improving the level of expression.

The above-described expression systems may be prepared using methods well known in the art. For example, various ligation techniques employing conventional linkers, restriction sites etc. may be used to good effect. Preferably, the expression systems of this invention are prepared as part of larger plasmids Such preparation allows the cloning and selection of the correct constructions in an efficient manner as is well known in the art. Most preferably, the expression systems of this invention are located between convenient restriction sites on the plasmid so that they can be easily isolated from the remaining plasmid sequences for incorporation into the desired mammal.

After such isolation and purification, the expression systems or constructs of this invention are added to the gene pool of the mammal which is to be transgenically altered. For example, one or several copies of the construct may be incorporated into the genome of a mammalian embryo by standard transgenic techniques.

One technique for transgenically altering a mammal is to microinject the construct into the pronuclei of the fertilized mammalian egg(s) to cause one or more copies of the construct to be retained in the cells of the developing mammal(s). Usually, at least 40% of the mammals developing from the injected eggs contain at least one copy of the cloned construct in somatic tissues and these "transgenic mammals" usually transmit the gene through the germ line to the next generation. The progeny of the transgenically manipulated embryos may be tested for the presence of the construct by Southern blot analysis of a segment of tissue. If one or more copies of the exogenous cloned construct remains stably integrated into the genome of such transgenic embryos, it is possible to establish permanent transgenic mammal lines carrying the transgenically added construct.

The litters of transgenically altered mammals may be assayed after birth for the incorporation of the construct into the genome of the offspring. Preferably, this assay is accomplished by hybridizing a probe corresponding to the DNA sequence coding for the desired recombinant protein product or a segment thereof onto chromosomal material from the progeny. Those mammalian progeny found to contain at least one copy of the construct in their genome are grown to maturity. The female species of these progeny will produce the desired protein in or along with their milk. Alternatively, the transgenic mammals may be bred to produce other transgenic progeny useful in producing the desired proteins in their milk.

EXAMPLES

EXAMPLE 1

Bovine Alpha S-1 Casein

We cloned bovine alpha S-1 casein with a cosmid library of calf thymus DNA in the cosmid vector HC79 (from Boehringer Mannheim) as described by B. Hohn

and J. Collins, *Gene*, 11, pp. 291–298 (1980). The thymus was obtained from a slaughterhouse and the DNA isolated by standard techniques well known in the art (T. Maniatis et al., "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory at p. 271 (1982)). We isolated the cosmid library using standard techniques (F. Grosveld et al., *Gene*, 13, pp. 227–231 (1981)). We partially digested the calf thymus DNA with Sau3A (New England Bio Labs) and ran it on a salt gradient to enrich for 30 to 40 kb fragments. The partially digested DNA fragments were then ligated with BamHI digested HC79 cosmid vector, followed by in vitro packaging by lambda extracts (Amersham) following the manufacturer's recommendation. The in vitro packaged material was then used to infect the *E.coli* K-12 strain HB101 followed by selection on LB plates containing 50 μg/ml of Ampicillin [Sigma].

We screened this library using a 45 base pair oligonucleotide probe: CAS-1. This CAS-1 sequence, 5' CATGGCTTGATCTTCAGTTGATTCACTC-CCAATATCCTTGCTCAG 3', was obtained from a partial cDNA sequence of alpha S-1 casein as described by I. M. Willis et al., DNA, 1, pp. 375–386 (1982). This sequence corresponds to amino acids 20–35 of mature bovine casein.

As a result of this screening, we isolated three cosmids (C9, D4 and E1). Partial subcloning of C9 and sequencing demonstrated that the cosmid represented a portion of the genomic sequence of the alpha S-1 casein gene.

We then synthesized several oligonucleotide probes corresponding to regions of the casein cDNA, based on published sequences [A. F. Stewart et al. *Nucleic Acids Res.*, 12, p. 3895 (1984); M. Nagao et al., *Agric. Biol. Chem.*, 48, pp. 1663–1667 (1984)]. Restriction mapping and Southern blot analysis [E. Southern, J. Mol. Biol., 98, p. 503 (1975)] established that cosmids D4 and E1 contained the structural gene and 9kb of upstream or 5' flanking sequences. The C9 cosmid contained the casein structural gene and 8kb of downstream or 3' sequences (see FIG. 1). We sequenced the cosmids E1 and D4 in the region corresponding to the transcriptional start of the casein structural sequence and determined that the sequence corresponded to that of the same region as described by L. L. Yu-Lee et al., *Nucleic Acid Res.*, 14, pp. 1883–1902 (1986).

We believe that the controlling region of Alpha S-1 casein is located upstream of the start of transcription. We have established after sequencing that there is a 40bp Exon I and that the signal sequence of CAS along with the sequences which encode the first two amino acids of mature CAS—arginine and glutamine —are found in Exon II.

We constructed the CAS promoter plasmid as follows: The genomic map of FIG. 1 shows that the control or promoter region along with Exons I and II may be cloned as a 9kb KpnI-BamHI fragment. Accordingly, we digested the E1 cosmid with KpnI and BamHI, then ligated it to pUC19 (Bethesda Research Labs) which had been previously cut with KpnI and BamHI. The resulting plasmid pCAS 1134 (see FIG. 1) contained the CAS promoter and signal sequence with a BamHI site suitable for cloning.

In order to allow the genomic construct to function in a eukaryotic host, i.e., to carry out transgenic work in which DNA is injected into the pronucleus, the prokaryotic sequences must first be removed. One method employed to remove prokaryotic sequences was to modify the pCAS 1134 so that the SalI sites flank the eukaryotic DNA. The KpnI site located upstream of the CAS promoter was changed to a SalI site using the CAS-11 linker 5' GGT CGA CCG TAC 3' which was ligated into the plasmid following digestion with KpnI. The resulting plasmid, pCAS 1141 (see FIG. 1) contained SalI sites flanking the CAS promoter and the BamHI cloning site.

EXAMPLE 2

Construction of the Cas-Recombinant Product Construct

One recombinant protein that can be produced by the process of this invention is tissue plasminogen activator or TPA. As demonstrated below, the casein signal peptide was used to direct secretion of TPA from the mammary glands of transgenic mice carrying a construct according to this invention. In this construct, the nucleotide sequence of the casein signal peptide was fused to the sequence of mature TPA by RNA processing. The sequence of TPA has been described in D. Pennica et al., Nature, 301, pp. 214–221 (1983). In the TPA gene, as in the CAS gene, there is a BamHI site in Intron II which separates the signal peptide from the mature sequence [R. Fisher et al., *J. Biol. Chem.*, 260, pp. 11223–1130 (1985)]. The cDNA of TPA shows the BglII site in Exon III at amino acid #3 of mature TPA.

We subcloned a 1.7kb fragment from the genomic clone of TPA [R. Fisher et al., supra] using BamHI-BglII. The 1.7kb fragment contained a portion of Intron II, the 3' splice acceptor site and Exon III up to the BglII site. This 1.7kb fragment was used to replace the TPA signal sequence found in the cDNA clone of TPA to provide a BamHI cassette. As shown in Example 1, there is a BamHI site located in Intron II which separates the sequence for the casein signal peptide from the sequence of the mature protein. The CAS promoter plasmid pCAS1141 was digested with BamHI and the BamHI cassette containing TPA was ligated into the digested plasmid, as shown in FIG. 1, to yield plasmid pCAS1151, which contains the CAS promoter upstream of the cDNA sequence of TPA. This construct allows the TPA structural sequences to accept the casein signal sequence by RNA processing.

We then isolated the DNA for use to transgenically alter mammals. We digested the pCAS1151 DNA to completion with SalI. Following electrophoresis in 1% agarose TBE [Maniatis et al., supra] the 13kb fragment corresponding to the eukaryotic sequences was cut out of the gel and the DNA electroeluted. We then centrifuged the DNA overnight in an equilibrium CsCl gradient. We removed the DNA band and dialyzed extensively against the buffer TNE (5 mM Tris, pH 7.4, 5 mM NaCl, and 0.1 mM EDTA, pH 8).

EXAMPLE 3

Transgenic Incorporation of the Construct Into Mice

The procedure for transgenic incorporation of the desired genetic information into the developing mouse embryo is established in the art [B. Hogan et al., "Manipulating The Mouse Embryo: A Laboratory Manual" Cold Spring Harbor Laboratory (1986)]. We used an F1 generation (Sloan Kettering) cross between C57B1 and CB6 (Jackson Laboratories). Six week old females were superovulated by injection of Gestile (pregnant mare serum) followed by human chorionic gonadotropin two days later. The treated females were bred with C57B1

stud males 24 hours later. The preimplantation fertilized embryos were removed within 12 hours following mating for microinjection with DNA and implantation into pseudopregnant females.

We injected the construct by first digesting the cumulus cells surrounding the egg with Hyaluronidase. The construct was injected into the pronucleus of the embryo until it swelled 30% to 50% in size. We then implanted the injected embryos (262) into the oviducts of pseudopregnant F1 females. Of 262 embryos injected and implanted, twenty three live pups were born. Tail blots of these were done and probed with nick translated pCAS1151 DNA, demonstrating that five of them contained the CAS sequence. Two of the female G0 progeny were cross bred to males at six weeks to produce a G1 generation. We tested the progeny of these matings for pCAS1151 sequences by tail blots. We then bred and milked the female obtained following parturition. Those female mice that carried the pCAS1151 DNA sequence produced TPA in their milk while the controls did not.

We mated transgenic male G0 mice with control females. We tested the G1 progeny by tail blotting and raised and bred for milking, females which carry the pCAS1152 sequence. The G1 progeny produced 0.2-0.5 μg/ml of TPA in their milk. We next crossed these females with a wild type F1. The progeny that carried the pCAS1151 DNA sequence produced the same TPA levels, while those that did not carry the sequence produced no TPA in their milk.

EXAMPLE 4

Transgenic Incorporation of the Human TPA Sequence Into Large Mammals

After at least one prior estrus period, sheep are superovulated before becoming embryo donors. More specifically, at about day 10 of the estrus cycle, each sheep is implanted with a progestagen-impregnated vaginal sponge (each sponge containing 60 mg 6 alpha-methyl-17 alpha-acetoxy progesterone). The sponge remains implanted for 12 days. Three days before the sponge is removed and until the day following removal, each animal receives a gonadotropin treatment, consisting of administration of 2.5 mg porcine follicle stimulating hormone by intramuscular injection twice daily. At the onset of estrus, the sheep are either hand mated to fertile rams or inseminated in utero with 0.2 ml per horn of washed ram semen. Within 72 hours of sponge removal, one cell fertilized embryos and cleaved embryos are surgically collected from the reproductive tracks of anesthetized sheep by retrograde flushing with about 6 ml Ham's F-10 medium containing 10% heat-inactivated fetal calf serum from the utero-tubal junction through the cannulated infundibular end of each oviduct. The flushings are collected and embryos removed under a dissecting microscope.

The embryos are then transferred to fresh Ham's F-10 containing 10% fetal calf serum and transferred to the stage of an inverted microscope equipped with micromanipulators. Each embryos is then microinjected with a plurality of a construct, such as pCAS to 1151, according to the process set forth in R. L. Brinster et al., *Cell*, 27, pp. 223-231 (1981). The embryos are then aspirated into a glass pipet tip with 10 ml Hams F-10 and expelled 1-3 cm into the fimbriated end of the oviduct in synchronized recipient sheep. These sheep then are permitted to gestate for the appropriate time and their progany are tested for incorporation of a DNA sequence coding for TPA. The female species of these transgenic offspring produce TPA in their milk.

A construct according to this invention containing plasmid pCAS 1151 is exemplified by a culture deposited in the American Type Culture Collection, Rockville, Md., on June 23, 1987 and there identified as LE392/pCAS1151, wherein pCAS1151 is in *E. coli* K12. It has been assigned accession number ATCC 67450.

While we have hereinbefore presented a number of embodiments of our invention, it is apparent that our basic construction may be altered to provide other embodiments which utilize the processes and compositions of this invention. Therefore, it will be appreciated that the scope of this invention is to be defined by the claims appended hereto, rather than the specific embodiments which have been presented hereinbefore by way of example.

We claim:

1. A process for the production and secretion into mammal's milk of an exogenous recombinate protein comprising the steps of:
   a. producing milk in a transgenic mammal characterized by an expression system comprising a casein promoter operatively linked to an exogenous DNA sequence coding for the recombinant protein through a DNA sequence coding for a signal peptide effective in secreting and maturing the recombinant protein in mammary tissue;
   b. collecting the milk; and
   c. isolating the exogenous recombinant protein from the milk.

2. The process according to claim 1, wherein said expression system also includes a 3' untranslated region downstream of the DNA sequence coding for the recombinant protein.

3. The process according to claim 1, wherein said expression system also includes a 5' untranslated region between said promoter and the DNA sequence coding for the signal peptide.

* * * * *